The Journal *of*
Parapsychology

Volume 10 MARCH, 1946 Number 1

EDITORIAL

TELEPATHY: WILL HISTORY REPEAT ITSELF?

THE TELEPATHY HYPOTHESIS is in a doubtful position. In our September number, telepathy and clairvoyance were reconsidered in the light of the experimental evidence available. It was pointed out there that, surprising as it may be, there is no experimental evidence of telepathy that cannot also be explained in terms of precognitive clairvoyance. The fact that, on the other hand, there *is* evidence which can be explained only by clairvoyance gives secure standing to that hypothesis, and leaves telepathy "up in the air."

This crisis, however, is not the first one for telepathy. It is almost the duplicate of a former situation that arose in the early thirties when the investigation of ESP first began at Duke University. When the problem of telepathy was taken up here, it was recognized with some amazement that *no true test of telepathy had ever been carried out up to that time.* There had been scores of so-called telepathy experiments, but they were as much clairvoyance tests as telepathy. No effort had been made in any of them to eliminate the possibility of clairvoyance. In the procedure followed in these tests, a percipient attempted to identify without the use of the senses some object that was being observed at the time by another person (the agent). The possibility that the object itself could be the stimulus for the percipient if clairvoyance were a possibility appears to have been considered negligible, for we have no record of a telepathy experiment conducted before 1930 without an objective stimulus.

It was not that clairvoyance was unknown. The neglect of clairvoyance as a counterhypothesis in telepathy tests could not have been due to a lack of evidence of that capacity. There had been experimental demonstrations of clairvoyance as far back as there had been studies of telepathy, although the former were far less numerous. We must suppose, rather, that telepathy was the much preferred hypothesis. Whatever the reason for this preference, it was so strong as to outweigh the experimental need for controlling telepathy tests against the possibility of clairvoyance, even though only a simple change in the procedure was required.

There is an important lesson in this bit of history. It will be well worth while to remember the puzzling fact that a half-century of research in telepathy was carried on without any challenge as to the validity of the method. We can only interpret this fact as indicating an undue degree of dominance of rationalistic over empirical criteria of truth in the parapsychology of that period.

When, in 1930, the question *was* finally raised, a new test was devised. But the innovation was a very simple one. The sender was instructed to select symbols in random order, without using any objective stimuli and without making any record until after the percipient had made and recorded his response. There was thus no object or objective record in existence on which clairvoyance could operate—not until after the trial was entirely over. Significant scoring occurred under these conditions, and this led to the conclusion that telepathy was, indeed, an actual occurrence. Clairvoyance was impossible in these "pure telepathy" (PT) tests.

Thus ended the *first* "reconsideration" of telepathy. Once the challenge was answered by the newly designed PT test, telepathy was given the status of an established phenomenon which it had been formerly believed to have. In *Extra-Sensory Perception,* which appeared in 1934, the hypothesis was submitted that telepathy and clairvoyance were simply modalities of a single basic function, extrasensory perception, differing only in the targets, thoughts, or objects upon which they operated. A fair amount of comparative observation supported this hypothesis, though it could only be regarded as suggestive.

Now, however, telepathy has a new competitor! With the estab-

General Library (P.R.R.)
Psychology
Direct
1-10-47
56149 (v.10, no.1-2)

lishment of precognition, the so-called pure telepathy test is no longer so pure as had been thought. This situation has already been presented in the article "Telepathy and Clairvoyance Reconsidered" referred to in the opening paragraph. It is sufficient here to say that once more, as in 1930, we find there is no experimental evidence which can be explained solely by the telepathy hypothesis. The combination of precognition and clairvoyance makes the delayed recording in the PT test hypothetically available to the percipient without requiring telepathy. The PT experiments therefore do not discriminate between telepathy and precognitive clairvoyance, and are therefore inconclusive.

Here, then, is where a glance at history may be instructive. It is evident from the discussion that has followed the appearance of the aforementioned article that there is still a powerful philosophical preference for the telepathic as against the clairvoyant interpretation of ESP. When we look back over the fifty-year period before 1930 and see how this intellectual bias could influence the researches of our predecessors, we can better realize how urgent it is for us to be sternly objective now in the solution of this present problem. We must be willing to take the answer only as it is indicated by the experimental findings. It is not a safe procedure to look first in the back of the book for it before we go to work on the problem.

The telepathy hypothesis, then, is up for re-examination. Again the older test procedures will have to be modified, this time to exclude both precognition and clairvoyance. It is not so simple to do that as it was merely to rule out the possibility of clairvoyance; but it can be done. Equipped with new tests, we shall find out in due time what we may rightly conclude concerning the occurrence of telepathy. Otherwise there is no great difference between this present issue concerning telepathy and the earlier one. In fact, the parallel is close. There is, of course, the minor point that in the earlier instance the new PT experiments were carried out before the problem was raised in publication. But today there are more co-workers in the field to help in the solution of the problem; and it is a problem upon which help is urgently needed.

But do we *now* have a valid test for telepathy? We must endeavor to make sure about it this time. It is clear today that had

we but thought more keenly fifteen years ago, we might have made our pure telepathy tests preserve their validity much longer; for we could have excluded precognition then as well as today. It was not an unheard-of concept at the time. It was just not taken seriously enough—exactly as clairvoyance itself was not taken seriously enough before 1930.

The greatest danger at present is probably from inertia. They who have been believing telepathy to be established tend to some degree to go on doing so in spite of the facts. When they are in such a state of mind, it will take too little to prove telepathy to them. Also they are not likely to conduct experiments themselves. They have "faith" that telepathy will survive its new ordeal and will be established again, more firmly.

Fortunately, the new tests are not extremely hard to conduct. As our scientific realism fully dawns upon us and we appreciate how much is at stake, there will be widespread experimentation, and the question about telepathy will in time be answered. If it is to be an affirmative answer, it may be an early one; if not, it will take a long time. But no one, we submit, can with any good ground of assurance, predict now *what* that answer will be. If he could, we would not need to experiment.

J. B. R.

THE PSYCHOKINETIC EFFECT: A REVIEW[1]

By J. B. RHINE

THE MIND exerts a causal influence over matter. Whatever we think of the nature of the mind or of matter, we recognize that some kind of interaction occurs between thought and the brain. It is a fact of common knowledge that conscious volitional experience initiates a chain of action that ends up in overt behavior—in measurable work if that be desired. Such action of the mind upon a physical system has long been listed in the dictionary as "psychokinesis."

Is psychokinesis, or PK, *confined* to the thought-brain relation, or can the mind directly influence a physical system outside the brain? This question has often been asked through the ages. It is the point of attack of this paper, and the facts that help to answer it are surveyed here. The survey therefore reviews the findings of experiments to test the effects of PK on external objective systems, not on the brain. Throughout the article from this point on, all reference to PK will be confined to its effect outside the organism.

Belief in PK is common in primitive societies. It is not unusual for men in preliterate groups to believe that the priest can exercise a direct influence over the physical world beyond the use of mechanical means. But the more literate, religious groups likewise have their doctrines regarding physical miracles wrought by powers attributed to spiritual (nonphysical) agencies over the material world. An occasional reference is made in the literature of social anthropology to physical effects produced by medicine men and fakirs. The suggested implication is left that these effects cannot always be accounted for in terms of recognized natural principles.

From the new branch of medicine, the psychosomatic, come reports of cures that appeared impossible but yesterday and are still unexplainable today. These are cases in which the state of mind of the subject is the determining factor in the organic adjustment

[1] There have been many requests for a general introductory account of the PK researches published to date. This article has been prepared to fill that need.

brought about. It is reported that warts are removed, blisters produced, hair falls out or turns gray, rashes come and go, all on the turn of the mind. Experiments involving the use of hypnosis have in some instances experimentally confirmed the therapeutic force of psychogenic factors.

Observations and experiments collected by the societies for psychical research have likewise made their contribution to the background of the problem of PK. There have been studies of strange, unexplainable physical effects associated with certain houses said to be haunted or with certain individuals said to be mediums. The studies of the mediums in some cases have gone far toward meeting the requirements for adequate experimentation.

So brief a sketch cannot do justice to the claims for PK. Certainly it would be difficult for those who are familiar with the branches of study involved to dismiss the evidence for PK coming from these sources as entirely worthless and unsound. Yet, on the other hand, it would be impossible to build a conclusive case for PK upon such evidence, even taken in its totality. There was not enough experimental control, enough repetition, enough observation, to justify the radical conclusion that PK can affect the environment. There had to be new experiments with better controls. It must be conceded, however, that the earlier observations played an important part in stimulating interest and leading up to research.

The ESP research, too, furnished a favorable background for PK. It may even be said that PK follows from ESP as a logical consequence. In the ESP of objects, we must suppose that if the object plays any part in the process—and how else can we think of it if we accept causality?—a psychophysical interaction is involved. If the interaction does not definitely involve the object itself, it involves some intermediating system such as radiation from the object, as happens in visual perception.

Such interaction leaves an effect upon the object or physical system. The effect may not be in any way registrable; we may have no instruments to discover it; but from all we know of interaction in nature, especially from the close parallel of sensory perception, we must infer that the object is acted upon. In beginning

research upon PK, then, after years of work with the ESP of objects, we were only looking for a way to measure the other end of an already familiar reaction.

The Dice Tests of PK and Their Advantages

The PK tests begun at Duke in 1934 were new in kind. They bore no resemblance to any of the above-mentioned earlier associations of PK. Not in religion, nor in the philosophies, nor even in the sciences, but in the world of gaming did we find a satisfactory method for an experimental tryout of the claim that the mind can influence matter. The game that furnished the idea is the commonplace and undistinguished one of crapshooting. The basis for a test of the PK hypothesis was furnished by the confidence which players often have that they can, when "hot," actually influence the dice to some extent to follow their desires without the use of trickery of any kind.

The dice test of PK has many advantages. The data from dice-throwing are among the oldest dealt with in mathematical statistics. Hence, the laws of chance applicable to dice are well known, and the special precautions are familiar. Dice-throwing lends itself very well, also, to the psychological needs of the PK test, since it provides an interesting, gamelike procedure. It is important, in dealing with any obscure, easily inhibited mental process like PK, to divert the subject or person being tested so as to keep his rational habits of thought from dominating him. A game, of course, makes one playful, somewhat relaxed, and more ready for anything new. It is a little world of its own that has its own rules. So the subject enters into it with a fresh outlook and a ready-for-anything attitude.

Such was the secret, we now think, of the success of the ESP card test introduced at Duke in 1930, and we followed the same principles as nearly as we could when the PK work began a few years later. In making a PK test-run consist of twenty-four single die-throws as a standard, we approximated the run of twenty-five trials in the ESP card tests. The simple scoring procedure made it possible for the subject to measure his success as he went from one run to another, and the interest in this scoring helped maintain his "morale" and preserve the gamelike character of the testing.

Controls of various types were possible, as we shall see, to insure against error in the interpretation of the results. In short, we had in dice-throwing the makings of an ideal test of the PK hypothesis.

Ideal tests do not, however, spring up spontaneously in exploratory undertakings such as this. We build them with many trials and not a little error; but of this early history, only the outline of developments can be given here. The full, original reports are to be found in this JOURNAL beginning with March, 1943, and from these a few salient features are described in this review.

PROCEDURES AND PROBLEMS

It was originally a young gambler who suggested dice-throwing as a basis for a PK test and proposed that we use "high dice" as the target (that is, the combination of faces we wished to have turn up when the pair of dice were thrown). "High dice" are eight or above. Accordingly, the first few series were made up of high-dice tests, but soon interest grew in comparing different numbers of dice per throw, and the target was changed to a given face of the die (such as ones, twos . . . sixes) fixed in advance of the run. At the start, the dice were thrown from the hand, but cups and mechanical devices were very soon introduced to eliminate the possibility of skilled manipulation. Finally, numerous controls were used to rule out the possibility that faulty dice might be the reason for any high scores obtained.

This problem of the defects or loading in the dice was the most important methodological issue, and it has been kept in the forefront throughout the PK researches. The first thought of the gambler or dice expert would be to suggest the use of "perfect dice," but to the experimenter the matter looks very different. For one thing, *no* die can be trusted to be perfect and remain so. For another, no manufacturer's guarantee is good enough for the purposes of these tests. We have to rely on our own experimental and statistical controls to be sure of our conclusions. With these controls, common, inexpensive, readily available, commercial dice are entirely satisfactory, and therefore they were used throughout the research except for experiments designed to compare special materials and variations in shape of the dice.

A number of experimental controls were used to take care of the question of faulty dice. For example, the subject threw the dice in certain series with no conscious effort to influence them. Again, series of PK tests were conducted in which the dice were thrown with equal frequency for opposite faces or combinations, or for all six faces of the die. Thus any effect of inequalities would be cancelled out. Further controls involved the comparison of two test conditions in which the same dice were used, the one condition favoring the operation of PK and the other designed to interfere with it (for example, to distract the subject). Any significant experimental difference thus obtained could not be ascribed to the defects of the dice. But the very best controls on the faulty-dice hypothesis are those which have emerged quite incidentally from new analyses of the data made years after the experiments were finished. These consist of significant differences in scoring rate due to certain effects of position of the trial in the test sequence, such as in the run. These position effects will be described in a later section of this paper.

The First Experimental Series

The scores that the PK tests yielded were fairly consistently above the level expected from chance. In high-dice tests the run (12 throws of the pair, 24 single die-throws) would be expected to give five hits (a hit being eight or above). There were 901 runs in the first experimental series which was carried out here in 1934. These gave 446 hits above the total expected (4,951 as against 4,505). On the assumption of true dice, this is *a highly significant deviation,* giving a critical ratio (CR) of 8.69, which would be expected only once in over a million billion series of such length.

The statistical analysis thus safely disposes of the hypothesis of chance and justifies our going on to deal with other questions. The hypothesis of "skilled throwing" as an explanation of the results is also ruled out by the fact that the highest scoring section of the data, an independently significant section, was obtained from tests done with a wholly gravitational method of throwing of the dice, a procedure that allowed no room at all for skill or trickery in their manipulation.

The question regarding the possible role of faulty dice in this series was answered equally well, and in the following way: Most of the high-dice tests were done in short sequences of three runs, and there was a marked decline of scoring in this sequence. The first run averaged highest (6.09), the second run next (5.15), and the third lowest (5.05). The difference between the first and the second run (like that between the first and third) is quite significant. Its CR of 4.33 signifies odds of 10,000 to 1 that such a result did not occur by chance. This alone is enough to establish the case against chance. Since the *same dice* were used in the first as in the second and third runs, and since the dice themselves could not have changed between the two runs, the conclusion is that the extrachance differences in these PK tests could not have been due to the imperfections in the dice. On the other hand, such a decline in scoring in the run sequence *is* entirely consistent with other studies of position effects, not only in the PK work but in ESP and in other psychological researches in which "primary effects," as they are called, have been noted.

Other Experiments and Conditions

I have gone over this first experimental series in some detail, only because it was the opening gun of the PK work. It was quickly overshadowed by better, longer, and more completely controlled experiments. Assistants and colleagues joined in the PK research, not only here at Duke, but also in a few other places where it became known. (There was no publication on this topic, however, until 1943, nine years later.) In the twelve years that have passed since its initiation, the PK research has been confirmed by more than a score of collaborators whose experiments are being reported in this Journal. Fourteen PK researches have already been carried out outside of the Duke Laboratory, ten of them with significant and confirmatory results.

One of the earliest innovations in the PK test was the introduction of a simple trip device to allow the uniform and impartial force of gravity to throw the dice. When released, the dice would roll over an incline to bounce around on a padded surface below. (See Figure 1.) The PK effect worked quite as well with this

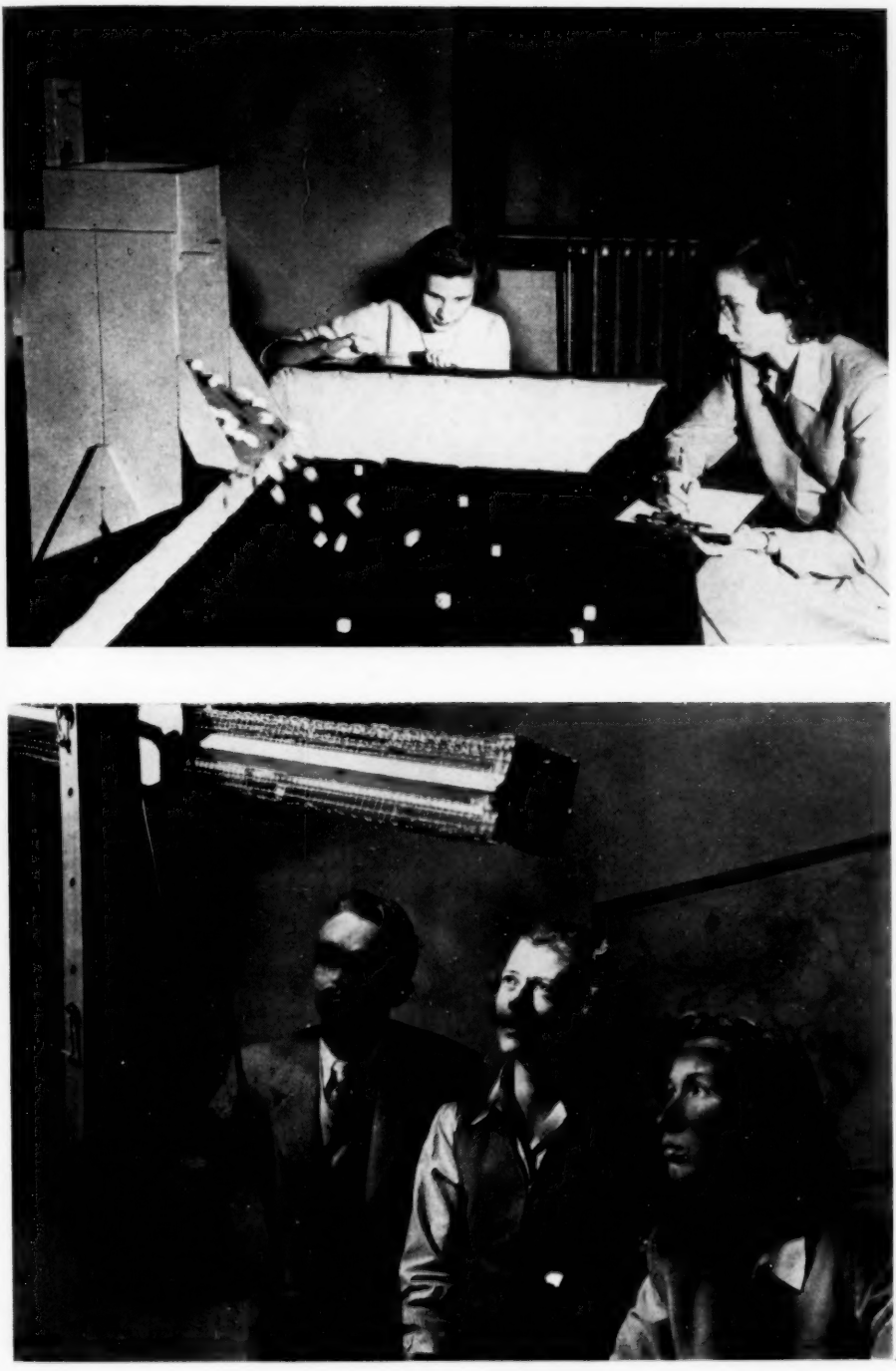


FIG. 1. The upper picture shows how the mechanical release of the dice allows them to fall by gravity upon a padded table. In the lower picture the dice are being thrown by the motor-driven rotating cage. They are read as the end of the cage passes the low point in its revolution.

"mechanical-release" method as with hand- and cup-throwing. Then we went on to the use of a long wire cage electrically rotated on a shaft through its midpoint, which let the dice bound from one end to the other. (See Figure 1.) The scores obtained with this method equalled the cup-throwing with the same dice and with the same conditions. The results showed that the subjects could mentally influence the mechanically rolled dice as well as those rolled from cup or hand.

Next in the series of progressive steps might be listed the throwing of an equal number of times for each face of the die. With this procedure, if any appreciable bias existed in the dice, it would cancel out in the general round-up of data. This method of control has been used in a number of experimental series, one of them the highest scoring series on record! In one of these around-the-die series, the electrical machine, too, was used; in the others, cup-throwing.

The evidence par excellence of the PK effect, however, derives not from mechanical nor from experimental controls, adequate though these have been. Rather, it is from the long, tedious statistical analyses that, years later, were carried out on the records—analyses that were quite independent of the original experiments themselves. I refer to the evidence from analyses for position effects.

Special Evidence of Position Effects

The ESP work had made us all conscious of the role of position of the trial in the run in determining scoring success. Declines of rate of scoring in the run (or in other units of data) had appeared in many ESP series, and we were consequently on the alert for similar declines in the PK records when, in 1942, they were subjected to general evaluation and analysis.

The very first PK work we did showed, as I have already mentioned, a decline of scoring from run to run in the set, or three-run sequence; that is, there was a *horizontal* decline on the record page. There was also a decline of success in the run or column; that is, a *vertical* decline on the page.

These discoveries of declines continued as we went from series to series of experimental records, and we found ourselves eventually

making analyses for position effects, both horizontal and vertical, in all the available records, even in research data that had been laid aside as deficient in some needed experimental control; for in examining records for distribution of hits on the record page, it mattered little whether, for example, the dice had been adequately controlled at the time. If, as was generally the case, there were significant differences between the scoring rate at the beginning (upper left) quarter of the record page and that at the end (lower right) quarter, it mattered not whether the dice were perfect. Such differences in scoring, if systematically recurrent and not due to chance, signify *psychical,* but certainly not *physical,* causation.

And these differences *were* recurrent and *were* extrachance! The analyses were concentrated on the quarter distribution (QD) of the page, which combined the horizontal and the vertical distributions. The record page was divided into quarters by a horizontal and a vertical axis, and the scoring rate was found for each quarter. The diagonal decline in scoring from the upper left to the lower right quarter was the natural resultant of the horizontal and the vertical declines and was, therefore, the primary point of interest. The diagonal decline was found very definitely to preponderate through the QD analyses. Of the eighteen series available (that is, suited to this analysis by reason of the character of the records) the QD of the page showed a diagonal decline in sixteen instances. With the eighteen series combined, the difference between the first and fourth quarters was highly significant.

The typical QD, then, was found to be one in which the first quarter was highest, the fourth lowest, with the other two in between. An example of such a QD is found by pooling the twelve series in which the dice were thrown for a *single* face as target (the Singles Series) instead of for a combination such as sevens or high dice. Figure 2 shows the total deviation from mean chance expectation for each quarter, the first being the upper left, the second the lower left, and so on. Note that the second and third are about halfway between the first and fourth. The CR of the difference between the total deviations of the first and fourth quarter is 5.56, and the odds are *a million to one* that chance would not give such a result.

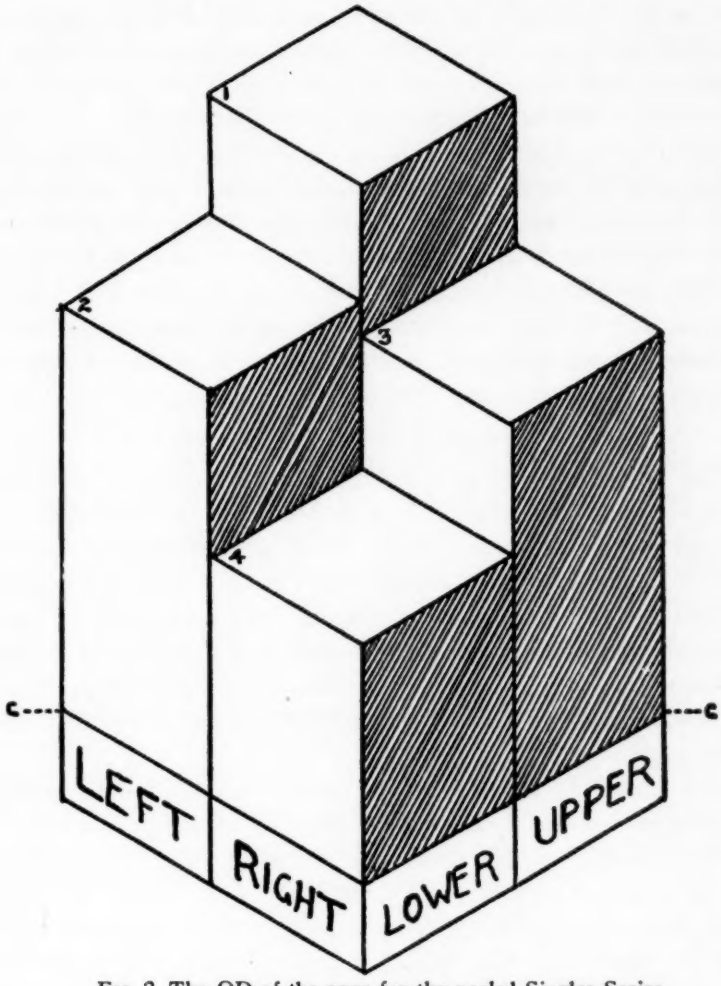


FIG. 2. The QD of the page for the pooled Singles Series in terms of the deviation of the average score for each quarter. The base line (marked "c") indicates chance expectation.

CONCLUSIVE NATURE OF THE EVIDENCE FROM POSITION EFFECTS

It will be apparent now, in considering the various counter-questions, why this QD evidence is indeed extraordinary. It is because, first, the difference in score distribution is beyond any reasonable explanation by chance. Second, this evidence rules out all

question of dice defects; the dice were the same throughout the course of the page record. Tricky throwing, recording errors, lost records—these and other hypotheses that might be proposed are all disposed of in unequivocal manner by the QD evidence; for not one of them could, with any reason, be supposed to have produced these significantly recurrent patterns of distribution of hits on the record page. It is important to keep in mind, too, that *no one participating in the experiment knew there would ever be an analysis for the QD.*

Thus the QD patterning was entirely incidental. PK was, as it were, writing its own strongest confirmation between the lines of the record sheets produced in these researches, evidence that came to light only with time and later analyses.

Accuracy of the QD Analyses

The QD's of the page, then, are quite important. There is all the more need in that case for us to determine the reliability of the analyses producing them. How "hard and irreducible" are their results? In answer to this question there are two lines of verification, different in effect, yet related. One is the confirmation of the *accuracy* of the findings through independently conducted analyses; and the other is the evidence of *consistency* in the findings, evidence obtained from the search for similar QD's in the subdivisions of the record page.

The answer to the question of accuracy is brief but conclusive. An independent analytical study of the QD of the page was carried out in 1943-1944 at the invitation of the Parapsychology Laboratory by a former member of the staff, Dr. J. G. Pratt. As a result of the reanalysis by Dr. Pratt the case for PK, as based on the QD evidence, was confirmed, and a report by him was published in March, 1944. An invitation was then extended to *any qualified committee or person* appointed by a scientific organization to repeat the QD analyses on the original records on file in the Laboratory at Duke. Any question that has thus far been raised, or seems likely to be raised, about the evidence for the PK effect can be pursued to its final answer in that way.

QD of the Set

The question of consistency of the QD brings us to a consideration of the smaller units of the record page. If, it was argued, there

is a general downward trend of scoring on the diagonal from upper left to lower right quarters of the record page, is it not probable that such a trend is evenly enough distributed to show up if we take smaller blocks of the page than quarters—eighths, twelfths, or sixteenths? Many of the series of PK tests, it so happened, were carried out in sets (that is, in small sequences of runs), with several sets to the page. The next step, then, was to do a QD analysis on the set similar to that done on the page as a whole. Nine of the eighteen series had sets suitable for this further distribution study.

The QD of the set was found to provide evidence of PK almost as reliable as that of the QD of the page. The declines of the two QD's are quite comparable even though there were only half as many series in the set analysis. The odds against chance are again of the order of a million to one. This indicates a higher rate of diagonal decline than that which the QD of the page showed for the larger group of eighteen series. The fact is that the nine series on which the set analysis could be made also gave a higher rate of diagonal decline in the page analysis than the rest, one quite comparable to that which they showed in the set analysis. So the consistency was really remarkable.

The main point here is a double one: In the first place, this analysis had to be conducted as an entirely independent one from beginning to end. That is, while the totals had to agree with those of the page-QD data, thus affording a check on the accuracy of both analyses, the entire set analysis had to be done by itself. Second, the results showed that the diagonal decline consistently extended down into the smaller units of the page—into the sets. We are thus insured further against any mere statistical artifact, as well as against faulty dice, skilled throwing, selection of evidence, recording errors, and the like.

QD of the Half-Set

The set analysis, however, does not exhaust this special type of evidence. The next step was to split each set in half and thus secure a still smaller section of the record page for analysis for the diagonal decline so evident in page and set. Moreover, there had been some degree of overlap (or "interdependence," statistically

speaking) between the whole set and the page, since most of the sets ran either the full length or width of the page and the two QD's thus shared a common axis in the quartering. An analysis of the *half*-set, however, would afford a completely independent statistical check. And to go a step further into the question of independence, this particular QD analysis was conducted by Dr. Pratt, whereas the other two QD studies (those for the page and the set) had been supervised by Miss Betty Humphrey.

To cap the climax of evidence, the QD of the half-set, too, showed the same persistent diagonal decline; this falling-off in scoring rate to the right and downward on the page and in the set now appeared also in the half-set. Again it was highly significant, with odds of more than a hundred thousand to one against chance occurrence; and again we have complete assurance that not chance, faulty dice, skill in throwing, recording errors, nor incorrect statistics can account for the results. This is "triple-distilled" evidence, for the original findings were themselves strongly indicative of the PK hypothesis. The three QD analyses are therefore supplementary verifications of this earlier conclusion—proof extraordinary of PK!

Some Psychological Findings about PK

We have reached the point, then, where it is quite in order to take the PK effect as established fact and to go on to a consideration of how it operates and what it signifies. This will mean discussion of evidence that offers more than mere proof of the occurrence of the PK effect—evidence, rather, of its nature and relationships.

In view of the prominent role played by position effects in this research, it is especially interesting that the declines obtained are not at all psychologically surprising. On the contrary, they are distinctly "human" and understandable. They are quite in line with the position effects found in ESP research, and they agree with decline curves found in experiments in learning and memory.

Again, on another point, the PK research appears to fit in with general psychological expectation. It was found that deliberate distraction of the subject's attention by the experimenter lowered the score average in PK tests to a point below that expected from chance

alone. Narcotic and stimulant drugs, hypnotic suggestion, and other factors and conditions have been found to influence performance in ways that are readily interpretable *psychologically* and cannot at all be accounted for in other terms.

Findings from Comparisons of Physical Tests

The most revolutionary outcome of the PK researches derives from the tests with varied physical conditions. There is where their uniqueness comes in. For example, the results of comparing large and small dice and different numbers of dice per throw offer evidence of the striking peculiarity of PK. In the several researches in which two or more sizes of dice have been used, no consistent relation has been shown between rate of scoring and size of dice (and that is to say between size of dice and weight, too, since the materials of the dice are similar). Whereas one pair of dice, with two or more times the volume of another pair, would be expected to require the influence of more physical energy, the results show no such expected relation. Sometimes the smaller, sometimes the larger dice are favored. Neither stands out consistently.

Furthermore, when different numbers of dice per throw are compared, more often than not the larger number gives better score averages than the smaller. For example, two dice have always been better than one. Six dice usually have been better than two. In another research, success with twenty-four per throw was higher than with six or twelve. Results from several experiments with ninety-six dice per throw compare favorably with scores based on the use of small numbers of dice.

Experiments with different objects than dice are just coming into publication. Two reports of researches made with two-sided objects (coins, chips, etc.) have already appeared in print, one from Cambridge University and the other from Duke University. These flat objects appear to be subject to the PK effect in somewhat the same degree as the dice. One report of a study of the effect of distance on PK has come out, too, showing no such decline as physical law would lead us to expect.

Some Interpretative Remarks

These comparative results and the ones which have of necessity been omitted in so brief a review seem without exception to lead to

the interpretation of PK as not governed by physical law of the type familiar to science today. It is, of course, difficult to find or accept an explanatory hypothesis unless it is stated in terms of familiar principles. But it is pretty clear from the PK research, as for some time now it has been clear from the ESP studies, that a new framework of scientific explanation is needed for dealing with certain psychical processes. *The mind is not a wholly physical system.*

Whatever this means in the long run to the psychology of the future and to the many fields and practices dependent upon psychology, it gives us at least a new way of inquiring experimentally into the nature of personality, first of all into the fundamental problem of the thought-brain relation. The question as to where and how the mental life of a man ties in with the physical world of which he is a part is the number one issue of the science of the mind. Psychology is an insecure science until that question is firmly answered, for not until then can we be sure where personality belongs in the universe. Unless these researches have changed its status, this primary issue of psychology is still in the stage of "I believe" or "I would argue" or "My theory is . . . ." It will be a revolutionary date in psychological history when one can say "I know" with reference to the way in which the experiencing mind works with its own body.

Everyone's philosophy of life depends directly upon what he thinks of his own basic nature, his place, his powers, and his destiny in the scheme of things. Yet of these matters we have only hypotheses or faiths. We need not labor, then, to justify the active exploration of an issue so fundamental to our way of life and code of valuation. More and more, those who are searching for a way of thinking by which life, individual and group, can be reliably and intelligently guided are coming back to the query: What is man with respect to the solid world of sense and mechanics? And today, in a scientific epoch, this issue leads us inescapably to the inquiries of parapsychology, the psychology which has been overlooked.

These inquiries into PK and ESP may be, to the effective human relations we hope for in the future, what the once dubious dabblings of scientists with unusual phenomena were to the more established sciences of today—an unrecognized and unwelcomed prelude. In

any case, so sorry a world as ours cannot be too selective in its dire need to find out enough about man's nature to be able to engineer him eventually away from self-destruction.

Parapsychology Laboratory
Duke University
Durham, North Carolina

BIBLIOGRAPHY

1. AVERILL, R. L., and RHINE, J. B. The effect of alcohol upon performance in PK tests. *J. Parapsychol.*, 1945, **9**, 32-41.
2. GIBSON, E. P., GIBSON, L. H., and RHINE, J. B. A large series of PK tests. *J. Parapsychol.*, 1943, **7**, 228-37.
3. ———. The PK effect: mechanical throwing of three dice. *J. Parapsychol.*, 1944, **8**, 95-109.
4. GIBSON, E. P., and RHINE, J. B. The PK effect: III. Some introductory series. *J. Parapsychol.*, 1943, **7**, 118-34.
5. HERTER, C. J., and RHINE, J. B. An exploratory investigation of the PK effect. *J. Parapsychol.*, 1945, **9**, 17-25.
6. HILTON, H., JR., BAER, G., and RHINE J. B. A comparison of three sizes of dice in PK tests. *J. Parapsychol.*, 1943, **7**, 172-90.
7. HILTON, H., JR., and RHINE, J. B. A second comparison of three sizes of dice in PK tests. *J. Parapsychol.*, 1943, **7**, 191-206.
8. HUMPHREY, B. M., and RHINE, J. B. Position effects in the large Gibson series. *J. Parapsychol.*, 1943, **7**, 238-51.
9. ———. PK tests with two sizes of dice mechanically thrown. *J. Parapsychol.*, 1945, **9**, 124-32.
10. MCMAHAN, E. PK experiments with two-sided objects. *J. Parapsychol.*, 1945, **9**, 249-63.
11. NASH, C. B. PK tests of a large population. *J. Parapsychol.*, 1944, **8**, 304-10.
12. ———. Position effects in PK tests with twenty-four dice. *J. Parapsychol.*, 1946, **10**, 51-57.
13. PRATT, J. G. A reinvestigation of the quarter distribution of the (PK) page. *J. Parapsychol.*, 1944, **8**, 61-63.
14. PRICE, M. M., and RHINE, J. B. The subject-experimenter relation in the PK test. *J. Parapsychol.*, 1944, **8**, 177-86.
15. REEVES, M. P., and RHINE, J. B. The PK effect: II. A study in declines. *J. Parapsychol.*, 1943, **7**, 76-93.
16. ———. The PK effect: the first doubles experiment. *J. Parapsychol.*, 1945, **9**, 42-51.

17. RHINE, J. B. Dice thrown by cup and machine in PK tests. *J. Parapsychol.*, 1943, **7**, 207-17.
18. ———. The PK effect: early singles tests. *J. Parapsychol.*, 1944, **8**, 287-303.
19. ———. Early PK tests: Sevens and Low Dice Series. *J. Parapsychol.*, 1945, **9**, 106-15.
20. RHINE, J. B., and HUMPHREY, B. M. The PK effect: the McDougall One-Die Series. *J. Parapsychol.*, 1943, **7**, 252-63.
21. ———. The PK effect: special evidence from hit patterns. I. Quarter distributions of the page. *J. Parapsychol.*, 1944, **8**, 18-60.
22. ———. The PK effect: special evidence from hit patterns. II. Quarter distributions of the set. *J. Parapsychol.*, 1944, **8**, 254-71.
23. ———. PK tests with six, twelve, and twenty-four dice per throw. *J. Parapsychol.*, 1944, **8**, 139-57.
24. ———. The PK effect with sixty dice per throw. *J. Parapsychol.*, 1945, **9**, 203-18.
25. ———. Position effects in the Six-by-Six Series of PK tests. *J. Parapsychol.*, 1945, **9**, 296-302.
26. RHINE, J. B., HUMPHREY, B. M., and AVERILL, R. L. An exploratory experiment on the effect of caffeine upon performance in PK tests. *J. Parapsychol.*, 1945, **9**, 80-91.
27. RHINE, J. B., HUMPHREY, B. M., and PRATT, J. G. The PK effect: special evidence from hit patterns. III. Quarter distributions of the half-set. *J. Parapsychol.*, 1945, **9**, 150-68.
28. RHINE, J. B., and RHINE, L. E. The psychokinetic effect: I. The first experiment. *J. Parapsychol.*, 1943, **7**, 20-43.
29. THOULESS, R. H. Some experiments on PK effects in coin-spinning. *J. Parapsychol.*, 1945, **9**, 169-75.

GESP EXPERIMENTS WITH THE FREE RESPONSE METHOD

By C. E. STUART[1]

ABSTRACT: This report deals with two series of ESP experiments. An agent in one room looked at a drawing or picture. The respondent in another room directed his attention, in imagination, to the agent, and then made a drawing of whatever came to mind. Four such trials were usually made in an experimental session. The response drawings were later ranked with respect to the degree to which they appeared to be associated with the "stimulus" pictures. The subjects of the first series were college students "conscripted" from an elementary psychology class. The subjects of the second series were volunteers who came to the laboratory under the incentive of their own interest. Ranking of the responses provides a numerical score. The material was ranked by the agent, by the respondent, and by an independent judge. The first series yielded chance scores. The second series was scored significantly by the respondents. In this series the scores when the agent and respondent were closely related pairs (twins, married couples, and engaged couples) contrasted sharply with the scores when the agent and respondent were unrelated. Closely related pairs gave significantly positive results. Unrelated pairs gave significantly negative results. The negative scores of unrelated subjects occurred mainly in the first three sessions. Later sessions yielded chance results.—Ed.

IN AN ESP TEST with free material (that is, a test in which the range of stimulus objects is relatively unlimited and the respondent is accordingly free to respond with a wide range of ideas) there is no clear-cut line between a "hit" and a "miss." If a responding subject makes what is practically a facsimile drawing of the stimulus object, it is easily identified as a hit. If the response reproduces the essential idea of the stimulus, it is also, conventionally, a hit. But what about such cases as the following: stimulus: Soldier, and response: gun; or stimulus: Cat, and response: dog? These are not hits, but neither can the familiar associations involved in such responses be considered convincing misses.

In work with unselected subjects the number of clear-cut hits is so small that unwieldy masses of observations are necessary to get

[1] This research was carried out while the writer held the Thomas Welton Stanford Fellowship in Psychical Research at Stanford University. The report has been approved by the Psychical Research Committee and is Communication No. 15 from the Psychical Research Laboratory.

enough cases to provide reliable statistical material. The kinds of problems encountered are evident in the research of W. Carington (1).

In my research with the free-material kind of ESP test I have made an attempt to utilize hits of a far-fetched kind, and this method I have called the "free response method" in order to emphasize the idea that the greatest range of freedom of choice actually occurs not in the stimulus material itself but in the responses to it. In proposing the method I have been guided by two working principles (or hunches) about the way ESP achieves expression. The first of these principles is that we do not know whether the reproduction response (stimulus: Cow; response: cow) is more valid or typical of ESP than an association response (stimulus: Cow; response: barn). Accordingly, it is reasonable to set up a method whereby all kinds of recognizable relations between stimulus and response may be used to arrive at a score representing stimulus-response correspondence. The second principle is that the participants in the experiment, the agent, who uniquely understands his perception of the stimulus, and the respondent, who uniquely understands his response, are the judges most likely to recognize and select out valid relations.

These principles are applied by a technique called preferential matching. The stimuli of a given experiment are set out in random order before a judge, and the judge *ranks* them with respect to their association with, or similarity to, each response. Previous applications of this method gave rank scores that seemed in general to represent the "amount" of relationship apparent to common-sense judgment and which at the same time were fairly simple to compare with chance hypotheses (5).

The research here reported was planned to exploit the free response method as a method of eliciting extrasensory perception from unselected subjects. Two series of experimental sessions will be distinguished: Series S1, in which the stimuli were small, simple line drawings; and Series S3, in which the stimuli were more complex magazine pictures, usually projected onto a screen.[2] Each experimental session included two phases: the first, a period in

[2] Series S2 has already been reported. See Reference 3.

which the stimuli were presented and the responses drawn; the second, a period in which the subjects ranked the stimuli and responses to provide a score.

Series S1

Procedure

One participant, the agent, was stationed in a room three doors down the hall from the laboratory room where the respondent made his response. A buzzer-signal system connected the two rooms. While the agent looked at a stimulus drawing, the respondent made a free drawing. Usually there were four such trials in a session.

The subjects were students from the elementary psychology classes at Stanford. They were "conscripts" in the sense that there was a course obligation to participate as subjects in the various experiments of the laboratory. They signed up in pairs for appointments of about an hour. When they reported, they decided which should be agent and which respondent.

The stimulus drawings used were simple line drawings in ink or pencil on cards about the size of a bridge card. The drawings used were: Arrow, Flower, Candle, Rabbit, Toad, Gun, Head, Parking Sign, Formula, Book, Block, Star, "Kayo," a Psychological Diagram, and a Rubin Figure. Just before the arrival of the subjects for a session, I selected and randomized the stimuli to be used, choosing pictures as unlike in form and meaning as possible.

The agent was given the stimulus cards, usually four of them, in an envelope and was told to take one out and look at it each time he was signalled by the buzzer. He kept a record of the order. The respondent was given a blank sheet of paper and was told to visualize momentarily the fact that the agent in the room down the hall was looking at a picture of something, then to draw anything he wished on the response sheet. I gave the buzzer signals and retired to an office adjacent to the respondent's room during each response, a period usually of three to four minutes. I marked a code letter on each response as it was completed and kept a record of the order. After the responses were completed, the agent put his list of the order of the stimuli into an envelope and shuffled the stimulus cards. I shuffled the responses and took them to the agent's room.

The stimulus cards were now placed in a row on the table before the agent, and one of the responses was placed below them. The agent was told, "This is what —— drew while you were looking at one of those cards. Which stimulus card do you think you were looking at when he made this response? You may use any feature of similarity or association you like to justify your choice." After a first choice was made, the card chosen was set aside, and the agent was asked, "Suppose that was not the card after all. Suppose it was one of these three. Which one would you choose as being the most likely one now?" After the second choice, a third choice was elicited in the same manner. In this way the subject assigned a rank to each of the stimulus drawings in respect to that particular response. I recorded the order on a cross-sectioned record blank. For the succeeding responses it was usually unnecessary to repeat every step of the instructions for the ranking. The subject customarily ranked his four choices without even setting the stimulus cards aside.

The stimuli and responses were now returned to the respondent's room. The four responses were placed on the desk before the respondent with one of the stimulus drawings below them. He was told, "This is one of the drawings which —— was looking at when you were making your drawings. Which one of your responses do you think you made when he was looking at this card?" The rest of the matching procedure was the same as that for the agent.

After the respondent's ranks had been recorded, the agent was called in, and the lists showing the true order of the stimulus drawings and of the responses were compared. The true correspondences, the responses made during the time the agent was looking at each stimulus, could now be identified. Since all possible correspondences of the stimuli and responses had been given a rank on the matching record, the ranks given the true correspondences were selected out and circled.

Before they left, the subjects were asked not to reveal to others what stimulus drawings had been used.

The original plan of the experiment was to vary the number of responses systematically from three to six. This proved impractical

after trial, and a four-response session was adopted for the rest of the experiment.

Scoring

In each experimental session responses were preferentially matched by each participant. Later they were preferentially matched by an independent judge, Mrs. Frances Miller, who performed the same matching task as that of the respondent.[3]

The matching procedure assigns to every possible stimulus and response pair a *rank* from 1 to 4. These can be translated directly into *weights* from 4 to 1. A first choice has a weight of 4, a second choice a weight of 3, and so on. The *sum* of the weights attributed to the true correspondences (the four stimulus-response pairs that occurred in the session) constitutes the matching *score*. There are three scores for each session: the agent's score, the respondent's score, and the independent judge's score.

If there is no consistent relation between the stimuli and the responses intended for them, then the weight assigned by the judge to a given stimulus-response pair may be anything from 4 to 1, depending upon accidental associations. That is, the weight of any stimulus-response correspondence is equally likely to be 4, 3, 2, or 1, if chance relationships alone determine the matching. As is evident, the mean chance expectation of the sum of four such weights is 10. If an observed mean differs significantly from 10, then we may conclude that there has been a consistent, recognizable relation between the stimuli and the responses intended for them.[4]

Results

There were 45 experimental sessions. The responses of these sessions were ranked by the agent, the respondent, and an independent judge whose mean scores were 10.00, 9.73, and 10.49, respectively. The agent's mean score is exactly equal to the chance mean. The respondent's mean score is slightly negative, but not significantly so. The independent judge scored slightly better than chance,

[3] Two accidents of procedure exposed the true correspondences of two experiments to this judge before she completed the matching. These two sessions were rematched by a second judge, Miss Virginia Patterson, and the scores are incorporated here.

[4] The mathematics of the preferential matching problem has been presented by J. A. Greenwood. See Reference **2**.

but not significantly so. Table 1 lists these means, their standard deviations, and the *t*'s testing their significance (6, pp. 438 ff.).

Table 1

SCORES OF SERIES S1

Judge	Mean	SD	*t*
Agent	10.00	±2.73	.00
Respondent	9.73	±2.12	.80
Independent	10.49	±2.17	1.47
Chance	10.00	±2.24	

SERIES S3

Procedure

Series S3 incorporated a number of changes in the experimental situation. The stimuli were pictures from magazines such as *Life* and the *Saturday Evening Post* cut approximately six-by-six inches in size in order to be projected by a balopticon projector. The agent sat in a darkened room and, during each presentation of the stimulus, looked at the picture projected on a screen.[5]

The subjects were volunteers. Some of them were students who came to the laboratory for direction in research.[6] These students brought friends and acquaintances with them at other times. The subjects participated as often as they wished, and in many cases I acted as agent or respondent. No time limits were imposed upon any part of the session. Before the experimental period the experiment was discussed freely. The exploratory nature of the work was emphasized and the attempt was made to make the subject feel that his performance was important as a phenomenon, irrespective of its apparent success or failure. The social relationship between agent and respondent was in most cases merely that of friendship or acquaintance. In a few instances, however, closely related

[5] In a few instances a projector was not available, and the agent simply looked at the stimulus picture. There was no suggestion in the results that it made any difference.

[6] All of these people deserve personal credit for their participation, but I wish particularly to acknowledge my debt for the patient collaboration of Mr. Robert Anderson and Mr. David Hill.

pairs participated together. These will be given separate attention later.

As in Series S1, four stimuli that I had previously selected and randomized were given to the agent.[7] Various rooms in the laboratory were used, but there were always at least one wall and two closed doors between the agent and respondent. Four responses were made by the respondent. Stimuli and responses were recorded and shuffled, as before, in separate rooms. Then the agent and respondent (fully warned to make no remark or gesture that would give a cue as to the order) were brought together to rank the responses. The participants filled out their own matching forms. In Series S1 the agent matched each response in turn to the four stimuli arrayed before him, and the respondent matched each stimulus in turn to the four responses arrayed before him. In Series S3 both participants ranked the material in both of these ways.

After the matching was completed, the two records of true order were compared, and the matching records were scored for the true correspondences.

In addition all the material has been doubly matched by one of various independent judges.[8] Occasionally more than one independent judge ranked the data. The rule adopted here has been to use the chronologically first of these judges to provide the independent judge's scores. Further discussion of the scoring procedure is presented in an appendix on evaluation.

Results

The total number of experimental sessions was 112. These included the work of 31 respondents. Fourteen of these respondents served only once; six served in two experimental sessions; two, in three; one, in four; four, in five; and four, in six and more. The trials of each session were matched six times: twice each by the agent, the respondent, and an independent judge. The means, standard deviations, and *t*'s of the scores observed are given in Table 2.

[7] Except when I was the respondent. In this case the agent selected four pictures at will from a pile of magazines. These he shuffled carefully and presented in the resulting random order.

[8] In a few cases I matched the responses as an observer. The other judges have been Mrs. D. H. Pope and Miss Elizabeth McMahan of the Parapsychology Laboratory, and Miss Frances Schulze of the Psychology Department of Duke University.

Table 2

SCORES OF SERIES S3

Judge	Matching	Mean	SD	t
Agent	First	10.09	±2.36	.40
	Second	10.15	±2.50	.64
Respondent	First	9.81	±2.34	.85
	Second	9.64	±2.24	1.68
Independent	First	10.14	±2.33	.65
	Second	10.26	±2.37	1.15
Chance	.....	10.00	±2.24	....

The means are not significantly different from a chance expectation of 10.00 nor from one another. The agents' means and the independent judges' means are virtually identical. The respondents' second matching mean of 9.643 is so low, however, as to suggest some negative factor masked by the pooling of the scores.

Grouped Results. The totaling of all sessions obscures, however, the contribution of certain groups of experimental sessions that should be discriminated. Previous work had indicated that closely related pairs of subjects apparently did better than unrelated pairs; and in the original program of the research it was proposed to treat these groups separately.

The pooled scores also include sessions in which I was the respondent. While every effort was made to informalize the experimental situation, there is one respect in which my participation necessarily differed from that of the other subjects, a gross difference in experience and training. The other subjects were having their first experience with ESP experiments. Although there is no clearly established evidence of the effect of past experience, it would be unreasonable to suppose in advance that it is of no importance. The likelihood of a real difference is great enough to warrant the separation of the sessions with the experimenter as respondent into a special category.

Unrelated Subjects. With the closely related subjects and the experimenter excluded there are 84 sessions with naïve respondents in which the agent and respondent were not related. The scores

of this group are totaled on the sixth line of Table 3. The means of the agents' scores and the independent judges' scores are very close to the chance mean. But the scores of the respondents themselves are 9.536 and 9.202. The latter gives a t of 3.82, with P less than .0001, showing the score to be significantly different from the chance mean.

A further breakdown of the results into scores for first, second, third, and later sessions by the various subjects shows that the respondents' negative scoring occurs mainly in the first three sessions. The scores for the first session alone are significantly below chance, with t's of 2.89 and 3.40. The lowest score occurs in the third session, but it is not significant for the small number of cases. Later sessions give scores around the chance mean. In Figure 1 the mean weights are plotted with respect to the order of the experimental sessions. The graph presents essentially the same figures as the mean weight column in Table 3.

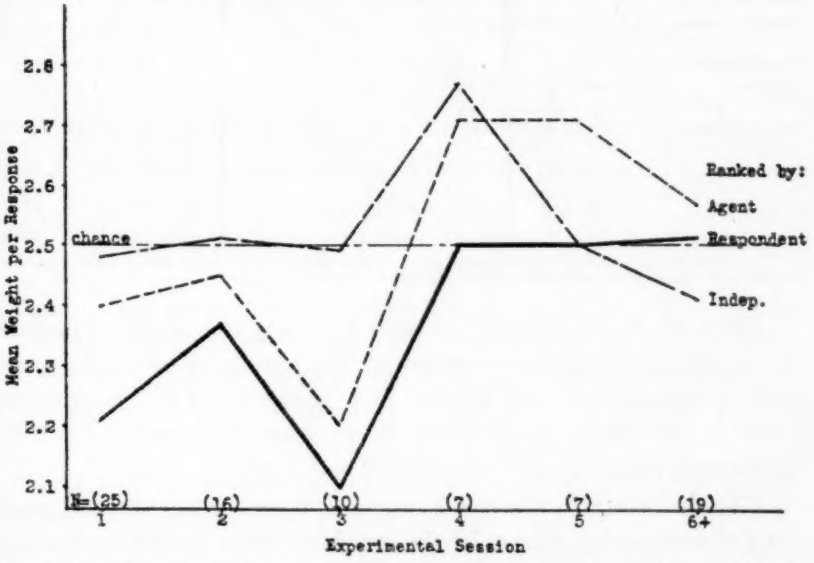


FIG. 1. Mean weights of responses of first and succeeding sessions with unrelated subjects.

Related and Unrelated Pairs. There were six sessions in which the agent and respondent were closely related. In two of these the

subjects were a pair of identical twins; in three they were married couples; and in the remaining one they were an engaged couple (who were married a few weeks later).

The responses of these sessions were given mean scores of 13.00 and 11.83 by the respondents, as may be seen in Table 3. These have *t*'s of 5.81 and 2.10, with associated probabilities of .001 and .05, and are thereby significantly different from a chance mean of 10.00.

Table 3

ANALYSIS OF THE RESULTS OF SERIES S3

Unrelated Subjects	Number of Sessions	Mean Weight per Response			Mean Score per Matching						*t* of Mean Scores					
		Agt.	Res.	Ind.	A_1	A_2	R_1	R_2	I_1	I_2	A_1	A_2	R_1	R_2	I_1	I_2
1st Session.....	25	2.40	2.21	2.48	9.64	9.56	8.92	8.72	9.88	10.00	.85	.84	2.89	3.40	.27	.00
2nd Session.....	16	2.45	2.37	2.51	9.75	9.88	9.75	9.19	10.13	9.94	.47	.27	.54	1.68	.19	.12
3rd Session.....	10	2.20	2.10	2.49	9.10	8.50	8.20	8.60	10.00	9.90	1.87	3.00	2.21	1.95	.00	.10
4th & 5th* Session	14	2.71	2.50	2.63	10.57	11.14	10.07	9.93	10.71	10.36	.81	1.75	.32	.16	1.22	.52
6th, etc.* Session	19	2.57	2.51	2.41	10.32	10.26	10.47	9.63	9.58	9.68	.64	.56	.85	.88	.70	.71
Total Unrelated Subjects.....	84	2.48	2.34	2.50	9.90	9.91	9.54	9.20	10.01	9.96	.40	.33	2.02	3.82	.05	.15
Related Subjects	6	2.90	3.10	2.92	11.83	11.33	13.00	11.83	11.50	11.83	3.05	1.40	5.81	2.10	1.96	1.70
Experimenter....	22	2.63	2.59	2.65	10.32	10.73	10.00	10.73	10.27	10.95	.49	1.07	.00	1.23	.55	1.52
Grand Total.....	112	2.53	2.43	2.55	10.09	10.15	9.81	9.64	10.14	10.26	.40	.64	.85	1.68	.65	1.15

*The fourth and fifth sessions, as well as the sixth and later sessions, were so grouped in order to provide a number of cases reasonably comparable with the first three sessions.

Since all the closely related pairs had only one session in which one or the other was the respondent, the sample of unrelated pairs most comparable is that of the first sessions. The respondents' scores in the latter, as is noted above were 8.92 and 8.72, both significantly negative, with *t*'s of 2.89 and 3.40.

The difference between the scores for the closely related pairs and the comparable sample of the unrelated pairs is also highly significant, with *t*'s equal to 5.04 and 3.55 for the first and second matchings, respectively.[9]

[9] The small number of cases of closely related pairs would seem to justify little confidence in this difference, but I present the matter fully here because additional observations clearly substantiating this one will be reported later.

Experimenter as Respondent. The pooled scores include 22 sessions in which I was the respondent. The results are itemized in the eighth row of Table 3. Although these sessions were given positive scores by all judges, none are significantly above a chance mean of 10.00.

Discussion

The results of Series S1 present little basis for comment. The mean scores are not significantly different from chance. No simple internal variation of a consistent kind is apparent. It is safest to conclude that this series illustrates a failure of the free response method to yield any easily assessable stimulus-response relationships.

The significant respondents' scores in Series S3 show a consistent, if negative, relation between stimulus and response. The usual counterhypotheses to ESP are rendered unlikely by the negative matching scores of the respondent and the fact that these were unique to him. In an appendix on evaluation at the end of this article, I have argued the case for the validity of the various scores. It is pointed out there that the agent's score is the one most likely to be spurious. The observed agents' scores are close to the expected mean of a chance distribution; so if they are assumed to be significantly spurious, they must be better than some "true" score, which would therefore have to be significantly negative.

If, in spite of the experimental precautions, cues to the true order were available to the respondent, they must have been used by him to avoid getting a good score. Yet, as we have noted, the initial attitude of the subjects was one of positive interest. The first session gave significant negative results. The contrast between the evident will of the subject and his score was easily observed. There was never any indication of a subject's being pleased by a poor score.

The variation in scoring from first to later sessions raises the question of selection. Did this not result simply because the poorer scorers dropped out and the more successful ones continued? There were seven subjects who served as respondent for five or more sessions. When these are selected out as a group, the mean weight scores per response of the group for the first five sessions respectively is 2.107, 2.286, 1.911, 2.679, and 2.500. The scores for this

group are slightly *lower* in the first three sessions than those of all the subjects. The observed variation, therefore, is clearly not a matter of subject selection.

From the standpoint of ESP, what sort of phenomenon has been demonstrated? The method is based upon the assumption that ESP will be evidenced by minor similarities and associations between stimuli and responses as well as by facsimile responses. The scores of the independent judges show that the amount of consistent objective relationship of this kind is apparently not significantly greater than chance. The respondents' scores are significantly different from chance, however; so there must be some consistent relation between the stimulus and the more comprehensive knowledge of the response available to the respondent. But since the scores are negative, the associations and similarities that guided their judgment must consistently have misled correct identification. This could have occurred in two ways. With reasonably complex stimuli the number of possibilities of relationship to justify matching are usually so great that the main problem for the judge is that of deciding which to use. The respondents may have systematically chosen the wrong association of the several which were possible. A second possibility is that the response may have avoided the stimulus systematically, but in a way recognizable to the respondent but not to the independent judge.

In either case the ESP phenomenon appears to be a consistent rejection or avoidance of the true stimulus-response correspondences by the respondent. In these experiments this phenomenon occurred in the initial experiences of these subjects with the ESP experimental situation. The first three sessions are significantly negative, while later sessions are about chance. It may be proposed that some learning or adjustment takes place, but whether the respondent learns to avoid the misleading relations or improves in the number of positive relations actually expressed in the response is not certain.

The significant positive scores of the closely related pairs, and the significant difference between these and comparable scores of unrelated pairs is statistically definite, but because of the small number of cases, must await later verification.

Appendix on Evaluation

Validity

The stimulus-response correspondences were matched by three persons: the agent, the respondent, and an independent judge. The scores of each present separate questions of validity. In the first place, if the assumptions of the free response method are tenable, the ESP response is recognizably related to the stimulus, at least to the degree that it is more related to it than other responses made to other stimuli. A relationship, however, may be valid and yet be so obscure as to be difficult to recognize.

The three judges have three different sets of experiences within which to select what they consider to be discriminating relationships. The agent has not only the stimuli but his associated experience while looking at each stimulus picture to compare with a limited number of expressions and meanings that the respondent has been able to convey in his drawings. The respondent has the advantage in that he is comparing the stimulus (which may be complex but is not obscure or uninterpretable, as a response may be) to a response which he is able to understand and interpret better than any observer. The independent judge has the task of educing relations between the stimulus picture and the response drawing without the benefit of personal interpretation of either.

Any prior decision as to which of these is the most valid discrimination of relations produced by ESP would have to be arbitrary; for excepting the unlikely case in which the independent scores would be significant and the scores of the participants negligible, differences in scoring by the judge might readily be attributable to differences in the way ESP influences the response in this particular experiment.

There is another question of validity, however. The agents' rankings are open to question because of two sources of cues which the agent may have used to guide his judgment of correct correspondences. When the respondent is permitted to take as much time as he wishes, he may in some cases make a simple response that takes little time, and in other cases draw a complex picture that takes a long time. The agent may note these time differences; and when he comes to rank the correspondences, he may rank them in terms of the amount of time apparently devoted to them, rather than in terms of association and similarity. A second source of cues is available when, for example, the agent knows the respondent well. He may be able to estimate which response idea was likely to have been the first or last, or to have had some recognizable serial relation to the others.

In one respect these possibilities of spurious agents' scores are relatively academic. The agents' scores were not different from chance;

so hypotheses to account for a difference are unnecessary. But if the respondents' scores are assumed as a true measure of existing stimulus-response relations, then the difference between the respondents' and agents' scores may well be accounted for by the possibility of timing or inferred seriality cues. The agent's score is therefore of questionable validity.

The respondent had no access to extraneous cues, except for one possibility. In Series S3 he matched the responses while the agent, who knew the correct order, was in the room. It would have been hypothetically possible for the agent to have connived with the respondent to use a simple code that would escape the experimenter's eye. Or, perhaps more likely, the agent might unconsciously have "given away" the order by some thoughtless word or gesture. The strongest evidence against this possibility is provided by the scores themselves. As noted before, the subjects gave every indication of being positively motivated and wanting to succeed. Yet they scored consistently below chance. There can be little doubt that the scores of the respondents represented their own judgment of perceived relations.

The independent judge's score is relatively free of question on counterhypotheses. The judge was given the randomized stimuli and responses in a folder and instructed to match them. In most cases he was unaware even of the source of the observations. The judge has one disadvantage that may be important: he does not share the motivation of the subjects. The participants' interest in their own work usually maintains incentive to study the possible relations between stimuli and responses with a degree of care that the most conscientious judge would find laborious.

The above argument is offered to support the view that since criteria are lacking as to what constitutes a good judge or a poor judge of ESP material of this sort, the most valid score is that of the respondent, who best understands his response.

Scores

Each judge matched the material two ways. In his first matching he ranked the four stimuli laid out before him with respect to each response. He then ranked the four responses arrayed before him with respect to each stimulus. The sum of the weights given the true correspondences constitutes the score in each matching. The two ways of matching are theoretically independent; but since they deal with the same material, the scores are correlated. This correlation is not a constant but varies with the complexity of the responses, the judge, and the breadth of criteria of relatedness the judge uses. When *all* the material is ranked by each of a set of judges, then the assumption of a relatively constant correlation is reasonable. This has been the

case in two previous articles (**3, 4**). But in the present work no judge has ranked all the material. Even the independent judges are various individuals.

To keep the results safely within the area of explicit mathematical assumptions the matching scores have been retained in separate distributions with a chance mean of 10. This results in the somewhat awkward circumstance that there are six scores for every observation. The distribution of each of these six scores is one to which the assumption of the Student *t* distribution may be applied for statistical evaluation.

Since each group of judges ranks the material twice, there are occasions when one score will be significant and the other not. In this case the criterion of significance has been whether twice the probability associated with the larger *t* is smaller than .01.

In referring to mean weight per response as ranked by a given judge, the average of both the judge's matchings has been used.

References

1. Carington, W. Experiments in the paranormal cognition of drawings. III. Steps in the development of a repeatable technique. *Proc. Amer. Soc. psych. Res.*, 1944, **24,** 1-107.
2. Greenwood, J. A. A preferential matching problem. *Psychometrika,* 1943, **8,** 185-91.
3. Stuart, C. E. A classroom ESP experiment with the free response method. *J. Parapsychol.*, 1945, **9,** 92-105.
4. ———. An ESP experiment with enclosed drawings. *J. Parapsychol.*, 1945, **9,** 278-95.
5. ———. An ESP test with drawings. *J. Parapsychol.*, 1942, **6,** 20-43.
6. Yule, G. U., and Kendall, M. G. *An Introduction to the Theory of Statistics.* London: Griffin & Co., 1937.

Parapsychology Laboratory
Duke University
Durham, North Carolina

A DIGEST AND DISCUSSION OF SOME COMMENTS ON "TELEPATHY AND CLAIRVOYANCE RECONSIDERED"

By J. B. Rhine

The manuscript "Telepathy and Clairvoyance Reconsidered," which appeared in the September number of this Journal, was sent in manuscript form to several friends, with whom I had been in previous correspondence on related topics, with the request that they read it and comment freely upon it. The replies contained much stimulating and valuable criticism, a great deal of welcome agreement, and further additions and improvements that strengthen some of the points made in the original manuscript. Each of the letters was in itself worthy of publication; but because there was a great deal of duplication among these letters, it seemed best to publish, instead, a digest of the principal comments, to offer a brief discussion of them, and finally to add a section on the several constructive suggestions contained in the letters. This procedure has been carried out with the approval of the correspondents. Their generous assistance is gratefully acknowledged. The commentators are Mr. Whately Carington, Mrs. K. M. Goldney, Dr. Gardner Murphy, Mr. Denys Parsons, Dr. S. G. Soal, Dr. Robert Thouless, Mr. G. N. M. Tyrrell, and Mr. Donald West.

I

Difficulties with Clairvoyance

The difficulties of making the clairvoyance hypothesis intelligible were discussed by Tyrrell, Carington, and West. Referring to my statement that "we are as ignorant of how to explain telepathy as we are of how to explain clairvoyance," West expressed disagreement, contending that "Carington's association theory, which he has recently developed into a comprehensive, idealistic philosophy, and Tyrrell's theory of apparitions are the first successful attempts to put psychical phenomena on a rational basis. Unfortunately, both

theories account for telepathy, but neither explains clairvoyance. Thus the a priori bias is toward telepathy and away from clairvoyance." Taking the other direction, Tyrrell also disagreed with my position. He argued: "As to fundamental explanations, it is indeed true that we have none for either telepathy or clairvoyance. But we know more of one factor in clairvoyance than in telepathy. We know more of material events and things than we do of thoughts or ideas."

Both Tyrrell and Carington find the explanation of clairvoyance so difficult as to be almost inconceivable. Although the two men take different philosophical positions, their difficulty with clairvoyance is much the same. Tyrrell concluded: "It seems to me: (1) that the proposition that a complete material thing with its sense-endowed properties, such as color and so forth, can be experienced clairvoyantly is not even intelligible; and (2) that while the proposition that the physical substratum of a material thing without its sense-endowed properties can be perceived clairvoyantly is intelligible (the sense-endowed property being supplied by the percipient), such a proposition would be as hard to establish as pure telepathy."

Tyrrell led up to this conclusion as follows: "If a subject clairvoyantly perceives a red disc, he perceives a quality of the disc which has been conferred upon it by *normal* vision. The disc per se is *not* red. . . . How is the disc known by the clairvoyant known to be circular? Does the clairvoyant perceive it from *every* point in space? This would give, not a material disc as we understand it, but a kind of Platonic *eidos* of a disc."

Carington wrote more emphatically: "No meaning whatever can be attached to the statement, 'This counter-in-the-dark is *now* red.' There is no conceivable means of verifying such statements. . . . You cannot possibly cognize the nonexistent; therefore you cannot possibly cognize, clairvoyantly or otherwise, the color of a counter-in-the-dark." He went on to say: "We must stick to the facts and not allow ourselves to confuse these with inferences." Referring to the clairvoyant sorting of invisible colored counters, he argued: "All you are strictly entitled to say is: 'At the start of the experiment the colored sensa were random, and at the end of it they were not.'

. . . I am in grave doubt as to whether an experimental demonstration of 'pure' and 'true' clairvoyance is even theoretically possible."

Tyrrell does not take this position of extreme positivism. He suggested that "because the cards used in a clairvoyance experiment have been seen and handled before the experiment and will be after, therefore the percipient is enabled to trace what happens to them when they are shuffled in the pack. Thus the apparent clairvoyance is really something more like telepathy."

West remarked in passing that "clairvoyance could be explained by a thorough-going idealism. . . . If nothing exists outside our ideas . . . the end result of a random card shuffle or dice throw is created by the mind of the onlooker." But he does not identify himself with this position.

Discussion

There is no question that the difficulties with the concept of clairvoyance are very great. Clairvoyance is impossible to understand in terms of our present knowledge. No combination of known principles can explain how twenty-five thin cards lying face-down in an enclosed deck a mile away can be clairvoyantly distinguished and perceived.

But this very inexplicability is what makes clairvoyance a *para*psychological phenomenon. We have no explanation for any parapsychical occurrences. It seems doubtful if it would make clairvoyance any easier to understand to try to account for it in terms of another unknown, telepathy; for, as Tyrrell says, we know still less about thought and ideas than we do of material objects. As a matter of fact, we cannot as yet even picture the essential difficulty involved in understanding telepathy.

In his effort at making telepathy intelligible, West made a suggestion quite in contrast with Tyrrell's: "If one accepts clairvoyance—that is, extrasensory perception of external objects—as a genuine faculty, there is no reason to suppose that human thoughts cannot act as external objects and be clairvoyantly perceived. Telepathy might be only a particular instance of clairvoyance." As long as it is a matter of speculation, then, a full hypothetical swing in either direction is possible; but we can *all* agree as to the difficulties

involved and *still concur in the need for adequately designed experiments* to determine what empirical findings are obtainable with respect to the problem of telepathy and clairvoyance. In the interpretation of the results we may do well, as Tyrrell suggested, to "take into account Professor C. D. Broad's analysis of the situation[1] before we can come to a clear understanding of what we mean by the term 'clairvoyance.' "

If one is impelled, as Tyrrell is, to extreme hesitation about accepting clairvoyance, he is naturally driven to lay down certain conditions to be met by clairvoyance experiments before they can satisfy his logical requirements. Tyrrell has been thinking along this line and has suggested in his Myers Memorial Lecture[2] a procedure for testing clairvoyance which would satisfy him if it gave significant results. By this method the subject tries to identify which one of five completely hidden boxes is lighted by an electric bulb. The box to be lighted is selected automatically, and records of the number of trials and hits are taken. There is no sensory experience of the boxes, for they are kept concealed. This last point is stressed by Tyrrell and is different from my suggestion of a clairvoyance test with his electrical machine.

Since, in my proposed clairvoyance test, the subject opens only the box he chooses in his response, I do not see how precognition can be of help to him. Even if we suppose perfect "teleprecognition" (to use Carington's coinage), I see no way for the box-opening to guide him; for he has to choose *before* the course of action is determined and thus before precognition of the consequence becomes a possibility. If it is already fixed, he has no true choice. We get into deep morasses here over problems of freedom and precognition; but as far as we can reason about it, Tyrrell's subject, if there was freedom of choice, opened up a course for precognition to follow only *after* her choice of box.[3] The choice

[1] Normal cognition, clairvoyance, and telepathy. *Proc. Soc. psych. Res.*, Lond., 1935, **43**, 397-438.

[2] *Apparitions.* London: The Society for Psychical Research, 1942. (See p. 90.)

[3] This paragraph raises the incidental but important question of whether there can be precognition of "potential futures"—possibilities that never become real events. West's remarks on this question appear in the last section of this digest. He holds that precognition is possible only for real future *events*. Thouless, in a later letter, states definitely: "Precognition can only be of what event will take place. If precognition could be a foreknowledge of the fact that if event A took

thus had to be a clairvoyant one. It should not, therefore, be necessary to conceal the boxes in order to eliminate precognition as a factor in producing the results. On the other hand, there is no reason for refusing to make the test if the apparatus is available and if the subject is duly prepared for it.

There was no sweeping criticism of the types of experimental evidence for pure clairvoyance as reviewed in the original article (that is, except for the logical difficulty over the concept of clairvoyance as such). We may therefore accept the fact of empirical clairvoyance with all its difficulties in the way we are accustomed to accept empirical facts in all of the sciences. As a rule, these difficulties evaporate with further exploration. If incontestable evidence is eventually obtained for telepathy, we shall accept it in the same way. It is doubtful, for that matter, if we shall *ever* obtain anything more inconceivable to us than precognition is today; but if scientific method has any value at all, that is no reason for rejecting the conclusion that precognition occurs.

II

Evidence for Telepathy from Mediumship and Spontaneous Parapsychical Occurrences

Tyrrell, Soal, and Goldney all believe that mediumship and spontaneous parapsychical experiences offer evidence of pure telepathy. Parsons and Murphy maintain that spontaneous experiences contribute evidence.

Tyrrell wrote: "I would add that observations of sensitives show that they can acquire a direct impression of people's moods and feelings, and this is evidence for telepathy." Soal stated that "a great many experiences with spiritualistic mediums are suggestive

---

place, then event X would follow, then the argument would break down and I don't see that any way would be left of demonstrating other kinds of paranormal cognition than precognition. It is generally assumed that precognition must be limited to foreknowledge of the event, but it might be as well to make this assumption explicit." I think we must all agree with that position. But in certain cases of spontaneous precognition, we may be dealing with other thought processes *combined* with precognition. This combination would be as plausible as that of reasoning and telescopic observations; and it could very well extend our thinking into potential-but-unreal futures, doubtless to our advantage. It would allow for an impression of a future event to be experienced, yet permit alteration of the course of action leading up to it. It would not be true precognition. The matter will come up more appropriately in later discussions of precognition.

of telepathy from the sitter's subconscious mind as their most natural explanation," and referred to his own case of John Ferguson[4] and to similar cases recorded by Osty, Hodgson, and others.

Goldney said: "I think if we take into consideration apparitions and mediumistic trance, we shall find the evidence for demonstrated telepathy increasing. . . . It is more difficult to obtain from a medium matters unknown to the sitter, which is understandable if telepathy is at work, but curious if, in fact, the medium is obtaining all her material by clairvoyance. (I leave out of consideration here other reasons to account for this fact.) Again, if telepathy is to be ruled out and clairvoyance alone postulated, it will be difficult to account for the effect noticed in Mrs. Piper's trance by Dr. Hodgson and reported in the *Proceedings of the S.P.R.* at the time; namely, that success varied with the *communicator* and was increased whenever the one communicator, George Pelham, held the field."

From Parsons came the suggestion that for both telepathy and clairvoyance the spontaneous cases have not been given "the consideration they deserve." "From some of these," he wrote, "I think a good case could be built up for telepathy."

Murphy, commenting on an earlier version of this manuscript, in which I took the position that certain spontaneous experiences favored telepathy, wrote as follows: "Personally I would strongly agree with you that spontaneous experiences support the idea of independent telepathy; in fact, I think they tend very strongly indeed in that direction." (See the letter to the editors from Dr. Murphy amplifying this point.)

Discussion

The preceding comments are for the most part general statements that were probably not intended to be conclusive on the question of telepathy. All of us are likely to agree that for so momentous an undertaking as that of proving the telepathy hypothesis, something more concrete and specific will be needed than we have thus far salvaged from the recent reassessments of the evidence. Perhaps such evidence exists. If it does and has been overlooked in the reconsideration of the question, it is important that it be

[4] A report on some communications received through Mrs. Blanche Cooper. *Proc. Soc. psych. Res.*, Lond., 1925, **35**, 471-595.

brought forward during the transition of values, concepts, and definitions which seems to be occurring in parapsychology.

Soal made the statement that in the case of John Ferguson, which he reported, his "private conjectures about a purely fictitious character were confirmed at the next sitting by Mrs. Blanche Cooper, the medium." "This happened so often," he said, "that I became normally certain that the medium was reading my subconscious mind." Soal's position is plausible according to our customary way of looking at things. But perhaps he kept notes on his thoughts in order that he could the more reliably check with the mediumistic utterances; and quite naturally when he obtained the record of the medium's statements, he recorded his thoughts then and eventually wrote a full report. All these are objective records, presumably accessible to clairvoyance. Telepathy would afford a more reasonable explanation if it were a demonstrated possibility. But as yet, it is not, while clairvoyance is.[5]

My colleague at Duke, the sociologist Dr. Hornell Hart, regards coincidence in dreams as evidence of telepathy. But since we do not know which person is the sender or original dreamer, we cannot be sure that the other dreamer (the "percipient") does not precognitively anticipate the oral or written account of the dream.

It has long been a generally accepted view, I think, that while spontaneous cases can have great value in stimulating and guiding the research, we cannot accept them as *final* evidence on any major issue because of the difficulty of controlling conditions, checking observations, and verifying memory and testimony. Likewise, while progress has been made in the methodology for studying mediumship, we have to take fully into account all the conditions under which a mediumistic study was made and what controls were main-

[5] Soal writes: "No written notes were made of my conjectures about 'John Ferguson' until after I had recorded the medium's confirmatory statement. If therefore these conjectures were accessible to clairvoyance, it could only be to a kind of 'precognitive clairvoyance.' Actually, the mental conjectures I made about 'John Ferguson' were so simple and so vividly present to my mind in the interval between successive sittings that I did not feel it necessary to make any record of them, especially as at the time I had not the remotest idea of any future publication. In January, 1921, I was investigating merely for my private satisfaction."

One can still inquire within reason: "How do we know that Dr. Soal himself was not guided in his thinking about John Ferguson by precognitive clairvoyance of what the medium was later going to state?" Thus we are running in circles. Experiment alone can end the argument.

tained before we accept evidence therefrom as conclusive proof of telepathy. It would be wholly unjustified to assert that these sources, mediumship and spontaneous parapsychical experiences, do *not* contribute evidence of telepathy; but it would also be hard to consider them as conclusive proof of that hypothesis.

III

In Defense of Telepathy in the Soal and Goldney Experiment

Soal, Goldney, and Parsons—all three have written in defense of telepathy as the best explanation of the results of the Soal and Goldney work with B.S.[6] Soal's words were: "In suggesting that B.S. failed in the clairvoyance experiments with certain agents because he was inhibited by his own mental preconceptions, I think you miss a very important point. When we changed the agent, the nature of the displacement underwent a radical change which I find hard to account for as a mere consequence of the subject's own beliefs. . . . I think it is true to say that B.S. had *no* pronounced theories of his own as to what ought to happen."

Goldney agreed: "I can only record my very strong impression that B.S. would have laid claim to both telepathic and clairvoyant faculty. . . . As regards our own bias, we were so pleased and surprised at getting marked positive results after so many years of seeming failure, and so intrigued at the unusual factor of seemingly persistent time-displacement into the future that I do not think we would have been unprepared to find a clairvoyance faculty to add to the other surprises. . . . Nor am I conscious of any bias against clairvoyance."

Parsons, like Soal, emphasized the change in the *nature* of the displacement effect with different agents and added: "I must put it on record that B.S. was equally confident in his ability to demonstrate pure clairvoyance with colored counters in my apparatus. (The results of both pure clairvoyance and telepathy experiments agreed with chance expectation.) With Soal and Goldney, B.S. failed in clairvoyance tests and succeeded in telepathy, even when

[6] Experiments in precognitive telepathy. *Proc. Soc. psych. Res.*, Lond., 1943, **47**, 21-150.

he did not know which of the two conditions would be imposed at a particular time. You argue that he might have known clairvoyantly which technique was being used, but is it not strange that he should get one hundred per cent success in detecting a change in technique, considering that his true cognitions of the card symbols never rose above thirty per cent?" Parsons then concluded that "telepathy is the *simplest* explanation of the Soal and Goldney results, just as clairvoyance is the simplest explanation of the screened touch-matching of Pratt and Woodruff."

Discussion

It is not easy to view this remarkable Soal and Goldney experiment with the degree of detachment which the status of the evidence for clairvoyance, precognition, and telepathy requires. If we decide only from the point of view of what seems reasonable, simple, and natural, from our customary ways of thinking in parapsychology, we shall certainly have no difficulty in favoring the decision for telepathy. Since we find, however, that we do not have evidence here (or elsewhere, thus far) that is *necessarily* interpretable as telepathic, and since we have found evidence from several sources for which there is no other interpretation than clairvoyance, the Soal and Goldney evidence cannot be interpreted as indisputable proof of telepathy. I do not see how we can either prove or disprove the telepathic hypothesis of ESP by means of this experiment. Telepathy is manifestly a clear possibility and is most plausible; if we had some adequate evidence for telepathy from other experiments, the telepathic explanation of B.S.'s performance would certainly be preferred. But as long as we can offer logically possible alternatives we cannot regard the experiment as definite proof of telepathy.

I am now convinced that B.S. would have welcomed evidence of clairvoyance, and accordingly I abandon the particular illustration I used to explain how he *might* have come, as he did, to fail to score in clairvoyance tests. But one can imagine several other possible ways to explain that fact. Some subtle factor in the way the card was being looked at, something in the personality of the sender who was looking at it, or some other one of a wide range of possi-

bilities in the physical situation surrounding the card may have operated unconsciously to help or hinder B.S.'s success. We have only to look at the fact that he could not produce for Parsons (with conditions allowing for telepathy) the successful scoring he did for Soal and Goldney to see that failure may be perfectly unexplainable. It is unexplainable, of course, because we know so little of what is really going on and cannot reason reliably about the effect of the changes in the situation. And since we cannot reason it out reliably, the design has to be in the experiment. There seems to be no way of closing in on this kind of problem except by definite experimental elimination of all the counterhypotheses. This is to state a platitude which no one has contested, yet which in this discussion needs to be kept to the fore.

Parsons compared the success with which B.S. seems to have avoided scoring above chance during the clairvoyance tests with his rate of success in the GESP tests (one hundred per cent as compared to thirty per cent). But there is another way of looking at this. We cannot rely on the assumption that B.S. was effectively aware of the clairvoyant condition because he failed in it. There may simply have been some essential condition lacking for success. This alone would give "one hundred per cent failure" (chance results) under clairvoyance test conditions as often as they were tried.

Perhaps B.S.'s ESP function required as a target situation the personal relation presented when someone was looking at the card. Let us suppose this was an unconscious requirement which had become habit with him. It is reminiscent of our first ESP experiments with Eileen Garrett. She began at once giving relatively high scores under conditions which involved a sender in an adjoining room thinking of a symbol, but at first gave only chance results in clairvoyance tests. Only after several thousand trials did she raise her clairvoyance scoring to a level that was evidential. Her explanation was that she was accustomed to working with people and liked the more personal arrangement.

We may agree, then, that Soal and Goldney obtained a demonstration of ESP that will be a classic, but there seems to be no way to tell with proper scientific finality whether the experiment involved telepathy or clairvoyance or both.

IV

Positive Contributions by the Commentators

Tyrrell inquired: "Can it be that the distinction between clairvoyance, telepathy, and precognition is unreal and is imported into the problem by the way in which our minds work?" While it may be objected that there is no point to discussing where telepathy will enter into the general picture of parapsychical abilities until it is experimentally established that the phenomenon actually occurs, Tyrrell's question presents a possibility which we might very profitably keep before us. The interrelations of the parapsychical abilities and their place in the total personality are of permanent interest.[7]

Parsons made the statement that "both clairvoyance and telepathy would seem to imply, first, a potential omniscience at the subconscious level, and, second, a selector mechanism which weeds out

[7] With reference to Tyrrell's question Thouless wrote: ". . . that is roughly the idea that I have been suggesting for some time. I think it would clarify thought if one restated the initial problem. Instead of asking, 'Can we get evidence for pure telepathy, pure clairvoyance, or pure precognition?' I should prefer to begin, 'We have got evidence of paranormal cognition (or psi). What are the conditions for it taking place?' Conditions to be investigated are: (i) the event has taken place; (ii) the event will in some future time be known to the S by normal perception; (iii) the event is present knowledge to some other person; (iv) the event will in some future time be known to some other person; (v) the event has not yet taken place but will take place in the future; (vi) the future event will be known to the subject by normal perception; (vii) the future event will be known to some other person; (viii) what is to be cognized is not an external event but exists only as the content of someone else's mind.

"The possibilities to be tested for possible determination of paranormal cognition by the subject seem to be:

(i) alone
(i) and (ii) together
(i) and (iii) together
(i) and (iv) together
(v) alone
(v) and (vi) together
(v) and (vii) together
(viii) alone

"These correspond to all possibilities I can see at present, but there may be others. Any of these conditions (or combinations of conditions) might be found to be necessary in the sense that, without it, paranormal cognitions would not take place at all. Alternatively, a number of them might be found to be sufficient conditions in the sense that with them alone psi cognitions took place although they might also take place with other combinations of conditions. Lastly, it might be found that certain of these conditions, although not necessary to paranormal cognition, were favorable to it in the sense that better scores occurred with them than without.

"So I should like to restate the problem as the experimental determination of what are the necessary, the sufficient, and the favoring conditions for paranormal cognition. That is, of course, merely a verbal restatement of the problem, but I think that put like that the issue is much clearer."

(more or less successfully) the irrelevant matter. In the case of clairvoyance, assuming telepathy is nonexistent, this selector mechanism is under the control of the subject alone." Where the selectivity and limitation come in in the total process of ESP represents a most important and interesting question, one which received the thoughtful attention of the late Mr. H. F. Saltmarsh.[8] It would be highly desirable to bring this "subconscious omniscience" hypothesis to test.

Parsons raised some interesting questions regarding telepathy and clairvoyance in animals, from which the thought derives that perhaps the experiments with telepathy in animals may be interpreted as more conclusively favoring telepathy than those with human beings. If we take Bechterev's experiments[9] with the circus dogs, or the tests which my wife and I conducted[10] with the filly, Lady, it is true we do not find the alternative of precognitive clairvoyance an acceptable possibility. The animal percipients cannot be supposed to have been able to interpret the objective records. Bechterev's dogs carried out complicated tasks, such as going to the next room and barking at a stuffed animal, picking up pieces of paper from the floor, and the like. In the tests with Lady, the horse responded by touching her nose to the appropriate one of a set of blocks on the table. (The target block was selected at random from a pack of small, numbered metal plates.)

On the other hand, even granted that there were no clairvoyantly perceptible cues (and we did, for example, stare at the block that was to be touched, though with our eyes screened and with motionless posture), there is still another counterhypothesis to be coped with in this animal work. We do not know where the ESP domain may end and where the PK effect may be considered applicable. It is just possible that the experimenter can direct the animal to some extent by some type of direct influence upon its organic system. With the present evidence of PK and the present lack of evidence of telepathy, it is not ridiculous to suspend judgment in the inter-

[8] The nature of extra-sensory perception. *J. Parapsychol.*, 1942, **6**, 101-10.

[9] *Zeitschrift f. Psycho-Therapie*, 1924, **8**, 280-304.

[10] An investigation of a "mind-reading" horse. *J. abnorm. soc. Psychol.*, 1929, **23**, 449-66.

Second report on Lady, the "mind-reading" horse. *J. abnorm. soc. Psychol.*, 1929, **24**, 287-92.

pretation of these results and to decline to accept them as conclusive evidence of telepathy.

West made the helpful suggestion that "there is one reason which I do not think you mention that makes screened touch-matching, in which the matching cards are not visible, difficult to explain on the precognition hypothesis. One could imagine the percipient saying to himself, 'If I put a card here, the checker will experience a success; but there, there, and there, he will experience a failure.' In other words, to make a choice [that is, by other means than clairvoyance of the card], the percipient must know all the different 'futures' attached to all the different choices." This reasoning does, in my judgment, strengthen the argument that the matching data support clairvoyance. It is similar to the logic used in the interpretation of the ESP-shuffle work, also to that used in the account of Tyrrell's clairvoyance tests with the electrical apparatus which was inserted in the manuscript after it was sent to the commentators. The general point back of the clairvoyance interpretation of all of these procedures is that the subject has to make his choice *before* his act makes *possible* a future consequence of it to be precognized. It does not seem as if we can profitably contemplate precognition of "futures" that are and always will be unreal. Thus the hypothesis of clairvoyance is necessary.

When Murphy suggested dowsing as a type of clairvoyance, a similar problem arose. His argument follows: "I agree that the matching experiments tend to favor clairvoyance. I would personally include the evidence of dowsing as of weight in favor of clairvoyance. See, for example, the book of Barrett and Besterman.[11] And since, in dowsing, the check-up follows the dowser's specifications—by digging at a particular spot and at a particular depth—it would be circular reasoning to say that the dowser makes contact with the experimenter's mind." That is to say that the dowser marks the spot where the well is to be dug before there is any future consequence possible.

The case is much the same as that of the percipient in the card-matching tests. The logic followed here is similar also to that of Parsons in interpreting the spontaneous parapsychological experi-

[11] *The Divining Rod.* London: Methuen, 1926.

ences he cited as evidence of clairvoyance. A clairvoyant dream may give rise to a definite course of action leading to the discovery of a body or a lost object. The fact that the *direction of action is necessarily decided before* a future course can be opened up for precognitive telepathy argues in favor of clairvoyance. This does not mean, however, that this evidence is on a par with the experimental work; but it has value, if only to indicate the scope of the function of clairvoyance.

One of the points of interest on which further comment will be most desirable is the proper procedure for testing the telepathy hypothesis. Thouless has proposed a valuable addition to the test initially suggested in "Telepathy and Clairvoyance Reconsidered." He proposes the use of a list of random numbers from one to five, each number to stand for a different symbol (assuming that the ESP card deck is used). At the beginning of the experiment, the experimenter must think of a number from zero to four which he is going to add on to the number taken from the list for a given trial. If the total of the two numbers is above five, five will be subtracted. The experimenter then thinks of the symbol which corresponds to the resulting sum (or difference).

This still leaves the case dependent upon the single observer. Thouless, however, commented: "The only objection I can see to this experiment is that it could not be witnessed, but I cannot see any way in which this could be overcome in a pure telepathy experiment, and I do not think it is important." Thus, until a better procedure can be invented, the pure telepathy test will have to depend, as many a scientific venture in the past has had to do, upon numbers of repetitions for its confirmation.[12]

I will try to summarize the situation as I now see it: The primary objective of the article "Telepathy and Clairvoyance Recon-

[12] Since this manuscript was first circulated, a further step toward devising a telepathy test has come to mind. The code number can be kept *subjective* and yet *shared* with an assistant by communicating it entirely in terms of common memories shared by the two—memories that never have had any objective expression or record. For example, if the number is five, the experimenter may say to the assistant, "The number is that which characterizes the size of family you will think about first when I mention carnations." The assistant may confirm the number with a reply in kind, but with different reference. This would go far toward restoring the telepathy test to as high a state of methodological objectivity as that of clairvoyance.

sidered" was to examine the experimental evidence for ESP to see whether it established the occurrence of either mode of extrasensory perception by the complete and final elimination of all alternative hypotheses, however difficult to conceive or far-fetched they might seem. Careful examination of all the quantitative experimental results showed that some of the results require clairvoyance to the exclusion of telepathy and all other variants of the ESP hypothesis, while no evidence for telepathy was found which could not at the same time be interpreted in terms of clairvoyance, either of present or future events. The discussions of the original article presented in this digest bring out many difficulties and show that the relative weight of the evidence for telepathy is strengthened by a consideration of spontaneous cases, but no evidence has yet been produced which absolutely requires telepathy as the sole explanation. Thus we may still say that clairvoyance is an *established fact* of parapsychology, while telepathy remains only a *logical possibility*.

Parapsychology Laboratory
Duke University
Durham, North Carolina

POSITION EFFECTS IN PK TESTS WITH TWENTY-FOUR DICE

By Carroll Blue Nash

ABSTRACT: Results obtained from nine subjects using 24 dice per roll gave a total number of hits which was highly significant, with a CR of 6.23. A positive deviation was obtained with each of the six die faces as target. When the results are tabulated in the chronological order of the roll, significant declines are found. Two distances, three and 30 feet, were compared.

The work here reported was done at American University where Dr. Nash was Assistant Professor of Biology; he is now head of the Department of Biology at Washington College, Chestertown, Maryland.—Ed.

This experiment, conducted in December, 1944, and January, 1945, was performed with 24 medium-sized dice, 15 mm. in diameter. Before each roll, the dice were placed in an open-topped container which was hinged to a three-foot long, inclined board covered with corrugated paper. A cord attached to the container, when pulled by the subject, overturned it, permitting the dice to roll down the inclined board into a cloth-lined, open-topped box in which the lower end of the inclined board rested.

When expressing a preference, the subject was permitted to select the target face, which was then kept the same throughout the set. The aim was to equalize the number of runs per target face, but because of preferences this was not quite achieved, the number of runs varying as follows: 112 for twos and fives; 128 for ones, fours, and sixes; and 144 for threes. Nine subjects, ranging in age from 15 to 44 years, were tested. Before taking the test they were asked if they believed psychokinesis is a scientific fact and if they believed they possessed the ability to demonstrate the PK effect. Six of the subjects replied yes, and the other three no, to both questions.

Each roll of the 24 dice constituted one run, and 16 runs constituted a set. The subjects rolled from three to 11 sets apiece, the

total data comprising 47 sets. Each set was divided into two half-sets of eight runs apiece. The subject's distance from the dice was alternated between three and 30 feet for each of the runs of the half-set. The distance used for the last run of the first half-set was also employed for the initial run of the second half-set, so that the distances for the initial runs of each of the two halves of a set were not the same. This change in the otherwise alternated distances between runs called the attention of the subject to the beginning of the second half of the set.

Table 1

TOTAL DEVIATIONS BY RUNS* IN THE SET

First Half-Set (Distance: 3 or 30 ft.)								
No. of Run...	1	3	5	7	2	4	6	8
Deviation....	+28	+56	+13	+ 3	+29	+34	+ 4	+ 7
Second Half-Set (Distance: 30 or 3 ft.)								
No. of Run...	9	11	13	15	10	12	14	16
Deviation....	+32	+38	+15	− 6	+ 3	+18	+20	+18

*Each item in the table represents a total of 47 runs with an expectation of 188.

Miss Mary Lou Courtney, a senior student at the American University, was the observer. She counted the number of dice that rested with the selected face uppermost, and the subject confirmed her count. The hits for each run were tallied on a record sheet similar to Table 1, and the subject was informed of the method of tabulation. Since the runs alternated between distances of three and 30 feet, they were not recorded chronologically across the page, but alternately on the left half and right half. The significance of this fact will be discussed later in this report.

It had been a part of the experimental design to compare the subjects' performance at two distances from the dice. About half the rolls were made from a distance of three feet, and the other

half at a distance of 30 feet from the dice. It was not realized, however, that the experimenter herself would have to be regarded as a potential subject, and she occupied a position near the dice throughout. Nevertheless, for what the comparison is worth, the results will be presented below for the two conditions.

Table 1 lists the deviations for each of the 16 runs in the set. The total number of hits obtained was 3,320, which produces a positive deviation of 312, an average run score of 4.42, and a CR of 6.23. The significant positive deviation cannot be attributed to irregularities of the dice and selection by the subjects of preferable die faces, as all of the faces showed a positive deviation from chance expectancy. (See Table 2.)

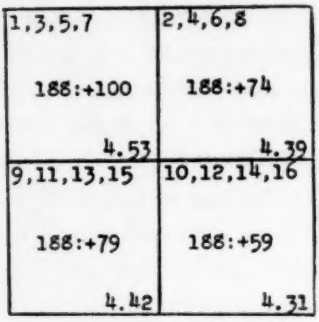
Table 2

DISTRIBUTION OF HITS FOR THE VARIOUS TARGET FACES

Target Face	Runs	Observed Hits	Expected Hits	Deviation	Run Score Average
1.....	128	573	512	+61	4.48
2.....	112	498	448	+50	4.45
3.....	144	603	576	+27	4.19
4.....	128	575	512	+63	4.49
5.....	112	489	448	+41	4.37
6.....	128	582	512	+70	4.55

The data are first analyzed according to position within the set as recorded. The following quarter distribution gives the hits for the combined sets. The squares show the chronological order of the runs, the total number of runs with the deviation, and the average run score for each quarter of the combined sets.

The quarter distribution follows the usual pattern of horizontal and vertical declines, as may be seen in the graphical representation of the QD. (See Fig. 1.) However, while the greatest difference between quarters is between the first and fourth, the CR of the difference (1.16) is not significant. The left quarters represent the left halves (but not the chronologically first halves) of the two half-sets, and the right quarters, the right halves of the two half-sets. (See Table 1.) The total deviation of the two left quarters is +179 and of the two right quarters, +133. The CR of the dif-



$CR_{d(1-4)} = 1.16$

ference is 0.92. Thus, the apparent decline between the first and second halves of the combined half-sets on the page lacks significance.

Variations within the combined sets cannot be attributed to differences in the subjects' distances from the dice, for, as will be shown below, the difference between the scores for the two distances is insignificant. Furthermore, initial distances are well balanced, with 26 sets beginning at the distance of three feet and 21 sets starting at the distance of 30 feet.

As has previously been indicated, the run scores were not originally recorded in their chronological order, but they are retabulated in this order in Table 3.

Table 3

TOTAL DEVIATIONS BY RUNS* IN CHRONOLOGICAL ORDER

	First Half-Set (Distance: 3 or 30 ft.)							
No. of Run...	1	2	3	4	5	6	7	8
Deviation....	+28	+29	+56	+34	+13	+ 4	+ 3	+ 7
	Second Half-Set (Distance: 30 or 3 ft.)							
No. of Run...	9	10	11	12	13	14	15	16
Deviation....	+32	+ 3	+38	+18	+15	+20	− 6	+18

*Each item in the table represents a total of 47 runs with an expectation of 188.

The quarter distribution of the combined sets with the runs arranged in their chronological order is shown in the QD square below. When the runs are arranged in this manner, the quarter

1-4 188:+147 4.78	5-8 188:+27 4.14
9-12 188:+91 4.48	13-16 188:+47 4.25

$CR_{d(1-4)} = 2.83$

distribution of the combined sets still shows a horizontal decline, which is pronounced between the two upper quarters. However, the vertical decline occurs only between the left quarters. (See Fig. 1.) The CR of the difference between the first and fourth quarters, which is 2.83, is now significant and is exceeded by the CR of the difference between the upper quarters, which is 3.39. The left quarters represent the first halves of the two half-sets, and the right quarters, the second halves of the two half-sets. (See Table 3.) The total deviation of the two left quarters is +238, and of the two right quarters, +74. The CR of the difference, which is 3.28, is now significant. Thus, when the runs are rearranged in their chronological order, there is a significant decline between the first and second halves of the combined half-sets.

If, as has been indicated by PK investigations, declines are typical of PK expression, the significant decline in the chronologically reconstructed set (in contrast to the insignificant decline in the set as actually recorded) indicates that the chronological order of the runs has considerably more effect upon the expression of PK than has the order of recording. This is due presumably to a greater degree of attention afforded the chronological order than the mere record itself.

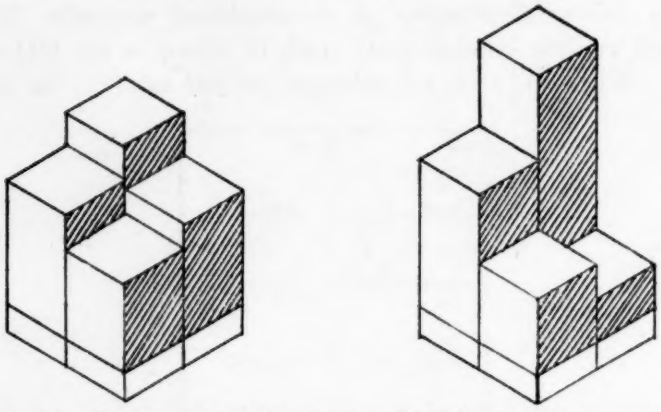


FIG. 1. On the left: quarter distribution of hits in the set as recorded. On the right: quarter distribution of hits in the set with the runs arranged in chronological sequence.

The distance of the subject from the dice had no significant effect upon the scoring. The total of the hits at three feet was 1,683, which yields a positive deviation of 179, a run score average of 4.48, and a CR of 5.05; the total of the hits at 30 feet was 1,637, which yields a positive deviation of 133, a run score average of 4.35, and a CR of 3.76. The CR of the difference between the two distances is 0.92, which is not significant.

The absence of a relationship between distance of the subject from the dice and degree of PK expression supports Rhine's hypothesis that the PK effect is extraphysical, and it is in agreement with the findings that fewer dice per roll (**1, 2, 3, 5, 6, 7**) and lighter dice (**2, 3, 4, 7**) do not favor extrachance scoring. If psychokinesis were accomplished by a physical force having its origin either in the subject or the dice, its effect would diminish with increasing distance between the subject and the dice.

The interpretation of the comparison of the two distances is qualified by the fact that the experimenter herself stood at the shorter distance throughout the entire series. It is possible that she may have influenced the dice. There is, however, some ground for rejecting this explanation: first, she was not consciously playing that role, whereas the subjects presumably were; second, the scoring rate varied widely, indicating that individual differences may

have been at work. For example, as noted below, the two oldest subjects gave results significantly different from the seven younger ones. But even though it is improbable that the experimenter's role was the determining one, the absence of a significant difference between the results from the two distances cannot be regarded as conclusive on the question of the effect of distance on the PK process. Yet it may be regarded as sufficiently suggestive to justify a further experiment.

The run score average was 4.43 for the six subjects who believed that psychokinesis is a scientific fact and that they possessed the ability to demonstrate the PK effect. While the run score average, 4.34, of the three subjects who replied in the negative to these two questions is lower, the difference is not significant.

The run score average of the two oldest subjects, ages 40 and 44, was 5.02, which was significantly higher than the run score average of 4.33 for the seven younger subjects whose ages ranged from 15 to 22. The CR of the difference is 2.99. This significant difference, however, is not necessarily an indication of greater PK expression in middle age than in youth, since the smallness of the sample may have admitted individual differences other than age.

References

1. Gibson, E. P., and Rhine, J. B. The PK effect: III. Some introductory series. *J. Parapsychol.*, 1943, **7**, 118-34.
2. Hilton, Jr., H.; Baer, G.; and Rhine, J. B. A comparison of three sizes of dice in PK tests. *J. Parapsychol.*, 1943, **7**, 172-90.
3. Hilton, Jr., H., and Rhine, J. B. A second comparison of three sizes of dice in PK tests. *J. Parapsychol.*, 1943, **7**, 191-206.
4. Humphrey, B. M., and Rhine, J. B. PK tests with two sizes of dice mechanically thrown. *J. Parapsychol.*, 1945, **9**, 124-32.
5. Nash, C. B. PK tests of a large population. *J. Parapsychol.*, 1944, **8**, 304-10.
6. Rhine, J. B. Dice thrown by cup and machine in PK tests. *J. Parapsychol.*, 1943, **7**, 207-17.
7. Rhine, J. B., and Humphrey, B. M. PK tests with six, twelve, and twenty-four dice per throw. *J. Parapsychol.*, 1944, **8**, 139-57.

Department of Biology
Washington College
Chestertown, Maryland

BOOK REVIEW

THE PERENNIAL PHILOSOPHY by Aldous Huxley. New York: Harper and Brothers, 1945. Pp. xii + 32. $3.00.

The Perennial Philosophy is an anthology of religious philosophy designed to demonstrate that the great religious leaders and teachers held virtually identical beliefs about the nature of the Reality behind the world of things and lives and minds. "This one Reality is such that it cannot be directly and immediately apprehended, except by those who have chosen to fulfil certain conditions, making themselves loving, pure in heart, and poor in spirit." It is as irrelevant to ask why this is so, says Huxley, as to ask why water should be composed of hydrogen and oxygen. The lives and works of the saints and sages constitute empirical establishment of the conditions and their fulfilment.

The first question essential to the Perennial Philosophy is the nature ot the true or divinely real Self. Who, in reality is the individual? The mystical answer is the dictum, "That art thou," embodying the principle that, first, the divine Self is the universe of all that is important beyond the human and mortal self, and, conversely, that all which is important in the universe is potential or immanent in the individual self.

That substantial aspect of divine Reality (God) attainable to human knowledge and will Huxley designates as the "Ground." "The divine Ground of all existence is a spiritual Absolute, ineffable in terms of discursive thought, but (in certain circumstances) susceptible of being directly experienced and realized by the human being." What, then, is the relation of the individual to the divine Ground?

The relation, says Huxley, is inexpressible, but analogous to that which holds between mind and body. "Mind affects its body in four ways—subconsciously, through that unbelievably subtle physiological intelligence . . . ; consciously, by deliberate acts of will; subconsciously again, by the reaction upon the physical organism of emotional states . . . ; and, either consciously or subconsciously, in

certain 'supernormal' manifestations. Outside the body matter can be influenced by the mind in two ways—first, by means of the body and, second, by a 'supernormal' process, recently studied under laboratory conditions and described as 'the PK effect.' "

With these two principles, Huxley guides us through what the Saints have said about more specific questions important to ethics, morals, and social living. "Personality," "Charity," "Good and Evil," "Silence," and "Idolatry" are a few of the many topics to which the Perennial Philosophy is applied.

To attempt to cite here a resulting list of rules of conduct for the saintly life would be to miss the point. Huxley is expounding a wholly mystical and nonrational philosophy, yet one which he feels is a practical guide to everyday ethics, morals, and religious belief. He wisely lets the Saints speak their own truths.

Many people interested in parapsychology are so interested because the findings seem to support basic religious beliefs such as those which Huxley outlines. Since inspired Truth is nonrational, it is perhaps pointless to expect that things which appear to support it do so rationally. But a certain congruence between the assumptions of the Perennial Philosophy and the findings of parapsychology is inescapable. The ultimate human values of the Perennial Philosphy are mental and spiritual ones, and the business of parapsychology is the direct scientific study of unusual capacities of the mind. Any sincere effort involving such similar purposes and values could hardly avoid similar consequences.

C. E. Stuart

LETTERS AND COMMENTS

August 16, 1945

Dear Dr. Rhine:

We have found it very difficult here to pick any holes in the brilliant series of PK reports which has come from your laboratory. At the same time, some of us here still have a residuum of possibly illogical but human skepticism which comes of not being able to reproduce the effects ourselves.[1] It may be that this very skepticism is inhibiting our results, but I certainly started my own experiments with considerable confidence after reading about the nine years of careful work which preceded the publication of your first paper.

Clairvoyance is not the only possible explanation of the PK results. In rotating-cage experiments, for instance, the subject could note the position of the dice at the start of the throw and, aided by an exact knowledge of all relevant elasticities and coefficients of friction, could compute the direction and strength of the PK force to be applied to each die. I grant that the gleaning of such knowledge amounts almost to a sixth sense.[2]

To me, decline effects are not an independent argument for the validity of PK and ESP,[3] but a disquieting argument against them. Consider a skilled factory operative engaged in a routine operation—say placing small pellets the right way up into the cavities of a mold. Study his output figures, note down the errors he makes, and plot these against Time on graphs. What is found? Both curves show well-marked terminal salience.[4]

Now put this same operative on to a very similar routine opera-

[1] *Proc. Soc. psych. Res.*, Lond., 1945, **47**, 277-300.

[2] Other possible but far-fetched explanations independent of clairvoyance occur to me. For example, the psychic force may have an inherent phototropic quality which can be altered at will by the subject. A "beam" of high phototropic value would make for the one-face because it contains the smallest area of black, while a photophobic beam would select the six-face. An additional assumption would be necessary to explain the Sevens Series, while the darkroom series (*J. Parapsychol.*, 1943, **7**, 124) remains unexplained.

[3] They are, of course, an independent argument for an extrachance effect.

[4] *Industrial Psychology* by Morris S. Viteles, Dept. of Psychology, University of Pennsylvania. Pp. 447, 525.

tion: instruct him to throw six dice so that they fall with a given face upwards. What is found? Again a well-marked terminal salience, identical in quality if not in quantity with the first one. *PK, whatever its nature, in fact behaves remarkably like skill.*

Herter and Rhine say: "Indeed, there is no known weakness of experimental procedure which would conceivably discriminate in favor of score success in the first half as against the second. No physical characteristic, no skill in throwing, and no weakness in the clerical work would produce a decline as the test advanced through the sequence of 36 throws for each target face, only to incline again in the last segment of six throws."[5]

I venture to disagree with this. Surely boredom or fatigue (working against skill) would produce just this effect. And if we admit PK, surely it is still fatigue which is producing the position effects. You may argue that subjects new to dice-throwing can hardly be classed as skilled operatives and that therefore, if a skill-like faculty was involved, you would expect scores to increase with practice and not expect the QD patterns actually found. To this I would reply that the skill-like faculty here exercised is evidently one with which the neophyte is already equipped; it does not have to be learned.

In rotating-cage experiments terminal salience again shows that some skill-like effort is at work, though manual dexterity is out of the question. There is a possibility that the subject scores high initially by using precognitive clairvoyance or telepathy to choose a favorable initial speed for the machine. My own rotating cage has given only chance results in 5,660 throws with 16 subjects.[6]

The manifestation of a skill-like faculty in experimental ESP and PK is surprising when we consider the evidence from spontaneous cases, automatism, trance, etc., all of which seem to show the psi faculty in quite another light, as an uncontrollable impulse often as its best when conscious striving is suppressed. For instance, the psychic powers of famous mediums such as Mrs. Piper and Mrs. Leonard did not noticeably fall off after the first five minutes; in fact some of their best material was given in the middle of sittings.

[5] *J. Parapsychol.*, 1945, **9**, 23.
[6] *Proc. Soc. psych. Res.*, Lond., 1945, **47**, 296-300.

With physical mediums, where one might reasonably expect a parallel with psychokinesis, the contrast is even more striking; it is usually stated that the "power" took considerable time to "build up," and then became progressively weaker towards the end of the séance.[7]

Another, and more important, line of attack on PK is raised by a letter from Charles E. Ozanne in the June JOURNAL.[8] He says: "It would be easy to construct a machine in which a delicately hung bit of metal or a slender wire rested near a metal plate in such a way that the slightest motion of the former would close an electric circuit." He explains how a number of people could sit around the instrument and strive to produce a movement, causing a current which would then be amplified to ring a bell. He goes on: "If by such a method no positive results were attained, grave doubt would be cast on the reality of the PK phenomenon."

Now instruments of the type described, usually known as "psychic galvanometers," have been made by the dozen ever since the beginnings of Spiritualism, and possibly before. Their inventors have usually claimed positive results. Many of those described in the literature were most carefully duplicated by Hereward Carrington in his laboratory and conscientiously tried out. Some of the instruments were extremely sensitive, many thousand times more sensitive, one would have thought, than dice-throwing. Yet not one of them yielded positive results under test. In many cases Carrington was able to point out the fallacy which had misled the inventor. The failure of the psychic galvanometers does not, of course, disqualify PK as an explanation of dice-throwing results, as there may be some quality in dice-throwing which is able to evoke the psychic force where delicately suspended wires fail. Nevertheless, Carrington's work is highly relevant to the problem of PK, and it has always surprised me that no reference has ever been made to it in the JOURNAL OF PARAPSYCHOLOGY. It is equally surprising that you have reported no work along these lines. The conditions in your laboratory would seem ideal for trying out some form of Alrutz's "volometer" or will-board.[9]

Yours sincerely,

DENYS PARSONS

[7] The evidence for the physical phenomena of mediumship is of course far from satisfactory.

[8] *J. Parapsychol.*, 1945, **9**, 138.

[9] *Op. cit.*, 169-85.

Comments on Mr. Parson's Letter

Thanks are due Mr. Parsons for his comments, which indicate a genuine interest in the PK experiments. There are a few points on which a word of clarification is needed, and I take the liberty of appending it here since it is highly important that readers such as Mr. Parsons who are also experimenters get as much of the research situation in mind as possible.

Resemblance of PK to Skill

Mr. Parsons argues that since fluctuations of accuracy in PK and in certain factory skills show similar curves, this is somehow against the case for PK. If we examine the reports more closely, however, the resemblance becomes only a superficial one. There are two special items of information that will help to clear up misconceptions about it.

First, the position effects which produce the curves in PK data to which Mr. Parsons refers appear in regular sequences or sets of trials, such as the column on the record sheet. But the curve is based not only on the *drop* in scoring in the column but equally on the *rise* occurring again in the next column, followed by another fall, and so on. This fact makes the suggestion of boredom and fatigue quite irrelevant. Moreover, in the case of the large Gibson series, for example, in which the set consisted of ten columns of only six entries each, this short column probably required less than a minute to complete. The comparison of the curve of this column (or even one that takes two or three minutes) with the fatigue curve of the day's work in the factory would seem too far-fetched to be of value.

Second, Mr. Parsons assumes that in the rotating-cage tests the subject varied the speed of the cage. He was allowed to set the speed at the outset, but there was no changing of rate made thereafter.

Reference to Other Efforts at PK Research

Mr. Parsons inquires why we have not made more reference to other experimental efforts along the lines of PK other than dice-throwing tests. There are *many* related studies that have never been adequately reviewed in our reports. A line has to be drawn some-

where, and wherever it is drawn, there is good prospect that someone will find it unsatisfactory. Spontaneous PK phenomena, studies of the physical phenomena of mediumship, psychosomatic effects, and other pertinent classes of manifestation might with profit be reviewed. A book, even a lengthy one, would be required; and we do hope such an inclusive review can sometime be written.

Use of Other Methods in PK Tests

Undoubtedly the diversification of methods will continue in the PK field. We have held to dice-throwing until recently (when some experimenters have taken up disc-throwing) because we wished to see first whether we had a thoroughly irrefragable case for PK on the basis of dice tests before branching out. Others are, of course, free to undertake other types of PK tests, and we hope they will do so. An important part of every research picture consists of how few man-hours there have been to devote to it.

J. B. R.

November 20, 1945

Dear Sirs:

In the lucid and comprehensive survey of the evidence for telepathy, clairvoyance, and various types of precognition presented in the September issue, I feel that a serious error of *omission* appears. This is the systematic discussion of *spontaneous* telepathy.

Psychical research came into being largely because of the great number of impressive cases of paranormal contact between persons separated in space, especially cases of awareness by one individual of what was occurring to another. Such spontaneous cases continue to be authenticated and reported. In the great majority of these cases, the *feelings* of the person undergoing an experience—for example, a catastrophe—are important. Indeed, the feeling tone is often the central point of the experience. Or, more broadly, the content of the spontaneous case is often something subjectively shared, not an objective event as such. Suppose we illustrate the general difficulty which appears when the clairvoyant hypothesis is

applied by summarizing a case appearing in *Phantasms of the Living*:[1]

The percipient, Mrs. Severn, reported that she woke with a start at seven in the morning, feeling that she had received a hard blow on the mouth, and that her upper lip was cut and bleeding. So strong was this impression that she held her handkerchief to her lip and was astonished to find that there was no blood. Later on, her husband came in with his handkerchief up to his lip. He explained that he had been out for an early morning sail, and that at about seven o'clock the tiller suddenly came around, striking him a bad blow on the mouth and cutting his upper lip, which bled a good deal.

In this well-known case of Mrs. Severn there is no trace of clairvoyant perception of the actual fact that her husband was boating on the lake nearby, and had at the time been hit on the mouth by the tiller; yet so intense was her feeling that she looked to see if her lip were bleeding. We seem here to be dealing with a case of paranormal sharing of an experience at a primitive sensory and emotional level, as we should expect in terms of telepathy; extrasensory contact with the *event itself* (as might be expected in terms of clairvoyance) is lacking.

A complete record of the apparent agent's experience, including its feeling tone, is seldom made at the time—a fact which necessitated the oblique, yet systematic types of corroboration used in the earlier English studies, especially a sort of cross-examination. Since there are no *records* of feelings to work from, such cases "make sense" as telepathy, not as clairvoyance. It is true that confirmation of spontaneous cases frequently relies in part upon diary notes, letters, etc., of agent, percipient, or both, or upon newspapers or other independent documentation. Usually, however, the percipient's experience goes far beyond such a nucleus of recorded fact, and is only verified by *later* conversation with the agent and with independent witnesses who remember what he has said.

It is true that the telepathic hypothesis cannot properly be stretched to include *all* spontaneous cases, for some of them are

[1] E. Gurney, F. W. H. Myers, and F. Podmore. London: Trübner & Co., 1886. See Vol. 1, pp. 188-89.

more simply interpreted in terms of clairvoyance—the percipient seems to see the event "from without," not from the agent's viewpoint. On the other hand, the majority of the well-authenticated cases require a good deal of stretching from the point of view of any hypothesis other than that of telepathy. Why not say, then, that the telepathic hypothesis, while strongly supported by many spontaneous cases, should not at present be regarded as the *sole* hypothesis worthy of consideration? And in planning further research, would it not be well to set up more experiments designed to put directly to the test the hypothesis suggested by the spontaneous cases?

It is true that we cannot insist upon telepathy with absolute certainty. But in all sciences, and especially in those less mature, we have to work with a limited number of probable, or at least plausible, hypotheses. The spontaneous cases seem to me to make both telepathy and clairvoyance plausible hypotheses, with the former considerably better supported.

Sincerely yours,

GARDNER MURPHY

GLOSSARY

In order to avoid constant redefining of commonly recurring terms in papers appearing in this JOURNAL, the following definitions are submitted for convenient reference.

*For a simple description of those terms marked by an asterisk, as they apply to the ESP test data, see Chapter VIII and the Appendix of *A Handbook for Testing Extra-Sensory Perception* by C. E. Stuart and J. G. Pratt. Further explanation may be found in any elementary statistical text.

AGENT: In tests for telepathy, the person whose mental states are to be apprehended by the percipient. In GESP tests, the person who looks at the stimulus object.

AVERAGE SCORE: Average number of hits per run.

CALL v.: To attempt to identify by ESP a stimulus object or a mental state of an agent.

CALL n.: The response described above; also the resulting selection.

*CHANCE: The complex of undefined causal factors irrelevant to the purpose at hand.

CHANCE EXPECTATION = MEAN CHANCE EXPECTATION: The most likely score if only chance obtains.

CHANCE AVERAGE: Mean chance expectation expressed as an expected score, generally in terms of average per run.

CHI-SQUARE: A sum of quantities each of which is a deviation squared divided by an expected value. Also a sum of the squares of CR's.

CLAIRVOYANCE: Extrasensory perception of objective events as distinguished from telepathic perception of the mental state of another person.

*CR (CRITICAL RATIO): A measure to determine whether or not the observed deviation is significantly greater than the expected random fluctuation about the average. The CR is obtained by dividing the observed deviation by the standard deviation. (The probability of a given CR may be obtained by consulting tables of the probability integral, such as Pearson's.)

CR OF THE DIFFERENCE: The observed difference between the average scores of two samples of data divided by the standard deviation of the difference.

DECK: Twenty-five ESP cards, five of each suit.

*DEVIATION: The amount an observed number of hits or an average score varies from the mean chance expectation or chance average. A deviation may be total (for a series of runs) or average (per run).

DIE THROW: The elementary unit of data in a dice-throwing test represented by the throwing and reading of a single die, no matter how many dice are thrown at a time.

EMPIRICAL CONTROL: An experiment which wholly or partially follows the main experiment with the exception that the conditions are designed to exclude the possibility of the hypothesis being tested.

ESP (EXTRASENSORY PERCEPTION): Response to an external event not presented to any known sense.

ESP CARDS: Cards, each bearing one of the following five symbols: star, circle, three parallel wavy lines (called "waves"), square, plus.

ESP SYMBOLS: See plate opposite page 1, this JOURNAL, Vol. 1, March, 1937.

EXPECTATION; see CHANCE.

EXTRACHANCE: Not due to chance alone.

FREE MATERIAL: The stimulus objects of experiments in which an unlimited or unspecified range of stimulus objects is employed (as contrasted with methods such as card-calling in which the subject knows that the stimulus object is one of a known range).

FREE RESPONSE METHOD: An ESP test method in which a relatively unlimited range of stimulus objects is used and in which the subject is instructed to express freely as his response any range of thoughts and ideas.

GESP (GENERAL EXTRASENSORY PERCEPTION): A technique designed to test the occurrence of extrasensory perception, permitting either telepathy or clairvoyance or both to operate.

HIGH-DICE TEST: A PK technique in which the aim of the subject is to try to influence a pair of dice to fall with the two upper faces totaling eight or more.

LOW-DICE TEST: A PK technique in which the aim of the subject is to try to influence a pair of dice to fall with the two upper faces totaling six or less.

MEAN CHANCE EXPECTATION; see CHANCE.

*P (PROBABILITY): A mathematical estimate of the expected relative frequency of a given event if chance alone were operative.

PARAPSYCHOLOGY: A division of psychology dealing with those psychical effects which appear not to fall within the scope of what is at present recognized law.

PERCIPIENT: The person who makes the calls in an ESP test.

PK (PSYCHOKINESIS): The direct influence exerted on a physical system by a subject without any known intermediate physical energy or instrumentation.

PRECOGNITION: Cognition of a future event which could not be known through rational inference.

PREFERENTIAL MATCHING: A method of scoring free responses. A judge ranks the stimulus objects (usually in sets of four) with respect to their similarity to, or association with, each response; and/or he ranks the responses with respect to their similarity to, or association with, each stimulus object.

QD (QUARTER DISTRIBUTION): The distribution of hits in the record page (or in a logical subdivision thereof, such as the set or the half-set) as found in the four equal quarters formed by dividing the selected unit horizontally and vertically.

RUN: A group of trials, usually the successive calling of a deck of 25 ESP cards or symbols. In PK tests, 24 single die throws regardless of the number of dice thrown at the same time.

SCORE: The number of hits made in one run.

TOTAL SCORE: Total of scores made in a given number of runs.

AVERAGE SCORE: Total score divided by number of runs.

*SD (STANDARD DEVIATION): Usually the theoretical root mean square of the deviations. It is obtained from the formula $\sqrt{npq}$, in which n is the number of single trials, p the probability of success per trial, and q the probability of failure. (For ESP cards, SD = $2\sqrt{\text{no. of runs.}}$) Sometimes the SD used is that estimated from the observed variability in the scores.

SD OF THE DIFFERENCE: For both ESP cards and PK tests using dice, the SD of the difference is equal to $\sigma_s \sqrt{1/R_1 + 1/R_2}$ where σ_s is the SD of a single run and R_1 and R_2 are the number of the runs in the respective samples compared. This gives the SD of the difference for run score averages.

SERIES: Several runs or experimental sessions that are grouped in accordance with a stated principle.

SESSION: A unit of an ESP or PK experiment comprising all the trials of one test occasion.

SET: A subdivision of the record page serving as a scoring unit for a consecutive group of trials, usually for the same target.

SEVENS TEST: A PK technique in which the aim of the subject is to try to influence a pair of dice to fall with the two upper faces totaling seven.

*SIGNIFICANCE: A numerical result is significant when it equals or surpasses some criterion of degree of chance improbability. Common criteria are: a probability value of .01 or less, or a deviation in the expected direction such that the critical ratio is 2.33 or greater.

SINGLES TEST: A PK technique in which the aim of the subject is to try to influence dice to fall with a specified face up.

STIMULUS OBJECT: The ESP card or drawing or other object, some identifying characteristic of which is to be apprehended by the percipient.

SUBJECT: The person who is experimented upon. In ESP tests, most commonly the percipient (though also the agent in GESP and telepathy). In PK tests, any individual whose task it is to influence the objects thrown.

TARGET: In ESP tests, the stimulus object; or in telepathy, the mental state of the agent. In PK tests, the faces of the die (or combination of faces) which the subject attempts to bring up in the act of throwing.

TARGET CARD: The card which the percipient is attempting to perceive (i.e., to identify or otherwise indicate a knowledge of).

TARGET DECK: The deck of cards the order of which the subject is attempting to identify.

TARGET FACE: The face on the die which the subject tries to turn up as a consequence of direct mental action.

TELEPATHY: Extrasensory perception of the mental activities of another person. It does not include the clairvoyant perception of objective events.

TRIAL: In ESP tests, a single attempt to identify a stimulus object. In PK tests, a single throw of the dice or other objects thrown.